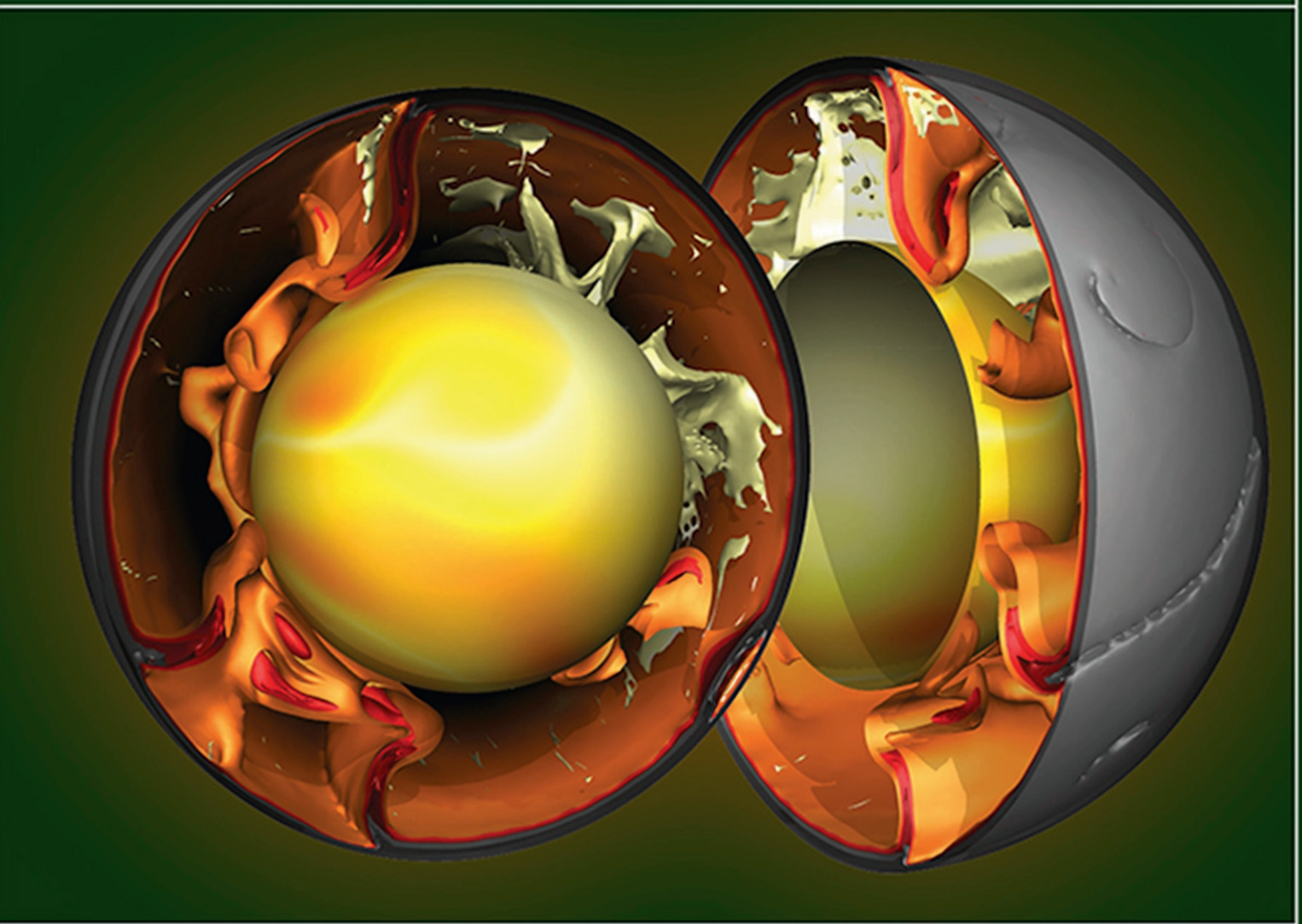
SECOND EDITION

TREATISE ON GEOPHYSICS

Volume 11

RESOURCES IN THE NEAR-SURFACE EARTH



VOLUME EDITOR: LEE SLATER

EDITOR-IN-CHIEF: GERALD SCHUBERT

TREATISE ON GEOPHYSICS

SECOND EDITION

This page intentionally left blank

TREATISE ON GEOPHYSICS

SECOND EDITION

EDITOR-IN-CHIEF

Gerald Schubert

Distinguished Professor Emeritus, Department of Earth, Planetary and Space Sciences, University of California, Los Angeles

VOLUME 11

RESOURCES IN THE NEAR-SURFACE EARTH

VOLUME EDITOR

Professor Lee Slater

Department of Earth & Environmental Sciences, Rutgers University Newark

ELSEVIER AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK
30 Corporate Drive, Suite 400, Burlington MA 01803, USA

Copyright © 2015 Elsevier B.V. unless otherwise stated. All rights reserved.

Cover image created for this work by Fabio Crameri, University College London.
The figure depicts the temperature field from a convection calculation using the code STAG-YY, written by Paul Tackley.

The following chapters are US Government works in the public domain and not subject to copyright:

1.11 Crust and Lithospheric Structure - Global Crustal Structure; 3.09 Earth Rotation Variations – Long Period; 4.07 The Role of Fault-Zone Drilling; 5.11 The Time-Averaged Field and Paleosecular Variation; 10.21 Exoplanetary Geophysics: An Emerging Discipline; 11.12 Tools and Techniques: Gravitational Method.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN (print): 978-0-444-53802-4

For information on all Elsevier publications
visit our website at http://store.elsevier.com/

Editorial & Production: Gemma Tomalin, Richard Berryman, Marise Willis, Michael Nicholls, Ellen Bishop
Cover design: Matthew Limbert

Printed and bound in the United Kingdom

15 16 17 18 19 10 9 8 7 6 5 4 3 2 1

CONTENTS

Preface *vii*

Editor-in-Chief *ix*

Contributors *xi*

Contents of all Volumes *xiii*

11.01 Resources in the Near-Surface Earth: An Introduction and Overview 1
L Slater

11.02 Physics of Porous Media: Fluid Flow Through Porous Media 19
Y Bernabé and A Maineult

11.03 Geophysical Properties of the Near Surface Earth: Seismic Properties 43
DR Schmitt

11.04 Geophysical Properties of the Near Surface Earth: Electrical Properties 89
PWJ Glover

11.05 Geophysical Properties of the Near-Surface Earth: Magnetic Properties 139
BM Moskowitz, M Jackson, and V Chandler

11.06 Tools and Techniques: Marine Seismic Methods 175
JOA Robertsson, R Laws, and E Kragh

11.07 Tools and Techniques: Ground-Penetrating Radar 209
J van der Kruk

11.08 Tools and Techniques: Electrical Methods 233
A Binley

11.09 Tools and Techniques: Self-Potential Methods 261
MD Jackson

11.10 Tools and Techniques: Active-Source Electromagnetic Methods 295
DV Fitterman

11.11 Tools and Techniques: Magnetic Methods of Exploration – Principles and Algorithms 335
Y Li and M Nabighian

11.12 Tools and Techniques: Gravitational Method 393
JD Phillips

11.13 Tools and Techniques: Nuclear Magnetic Resonance 419
M Müller-Petke and U Yaramanci

11.14 Tools and Techniques: Radiometric Methods 447
PG Killeen, CJ Mwenifumbo, and KL Ford

11.15 Mineral Resources 525
ME Best

11.16 Unconventional Fossil Fuel Reservoirs and Water Resources 557
E Barth-Naftilan and JE Saiers

11.17 Near-Surface Geophysics at the Hanford Nuclear Site, the United States 571
TC Johnson, DF Rucker, and DR Glaser

Index 597

PREFACE

The *Treatise on Geophysics* is a comprehensive and in-depth study of the physics of the Earth and of terrestrial-like planets residing both within and outside our solar system. Its breadth and detail of coverage are beyond what any single geophysics text can provide. The first edition of the *Treatise on Geophysics* was published in 2007, nearly a decade ago. Of course, there is much progress in science in that length of time and the field of geophysics has grown rapidly and developed new lines of inquiry. Accordingly, we have brought the *Treatise on Geophysics* up-to-date with this second edition. The new edition will continue to provide students and professionals with fundamental and state-of-the-art discussion of all aspects of geophysics. In this new edition, the reader will find updates to all the chapters contained in the first edition and new chapters that discuss topics missing from the first edition. In a few instances, chapters from the first edition have simply been reprinted in order to avoid gaps in coverage. Chapters that emphasized fundamental physics often required little updating.

A highlight of the second edition is a new volume on Resources in the Near-Surface Earth. The new volume discusses the role of geophysics in the exploitation and conservation of natural resources and the assessment of degradation of natural systems by pollution. The near surface is a zone where humans and natural systems interact. Understanding the effects of human impacts is a challenge particularly for the long term.

There are four new chapters in the volume on Deep Earth Seismology. One of them describes progress in seismic instrumentation in the last 30 years in the era of broadband digital sensor and recording systems. Another one addresses a new approach for the determination of crust and uppermost mantle structure, based on ambient noise cross-correlations. This has revolutionized the field, since it allows the study of regional structure at high resolution in areas where there are no earthquakes. The third new chapter is devoted to recent advances in characterizing the lithosphere–asthenosphere boundary under both oceans and continents. Finally, the last new chapter discusses the morphology of subducted slabs in the Earth's transition zone.

The Mineral Physics volume has new chapters on dynamic (shock-wave) compression methods, and on the mineralogy of the upper mantle and the D″ layer focusing on post-perovskite, and extrasolar super-Earths.

New material in the Geodesy volume expands the temporal resolution and latency of geodetic measurements to high-rate, real-time measurements and increases the spatial resolution to meter size pixels with optical and radar imaging systems. Application areas are extended with new material on the uses of Earth gravity measurements.

The Earthquake Seismology volume has a new chapter on rock mechanics that is a comprehensive account of modern rock mechanics with emphasis on the variation of frictional properties of rocks during an earthquake cycle. Many chapters have considerable updates reflecting the lessons from significant recent earthquakes (e.g., the 2011 Tohoku-Oki earthquake and the 2013 Okhotsk deep earthquake), new observations on slow slips and triggered earthquakes, and developments of new methods (e.g., back projection method). Some chapters are updated (e.g., paleoseismology and strong ground motion) with substantial new material.

Chapters in the Geomagnetism volume have been updated to include recent advances in the acquisition of larger data sets and higher quality data and in the construction of more reliable and sophisticated models. Improvements in data have occurred in areas such as magnetic anomalies, paleointensities, and planetary magnetism. Improvements in models of the present day field include so-called comprehensive models and models that incorporate data assimilation.

The Crustal and Lithosphere Dynamics volume has new contributions on mountain building and landscape evolution.

The Mantle Dynamics volume contains new chapters on the generation of plate tectonics from mantle convection and the core–mantle boundary region.

The Core Dynamics volume features a comprehensive update of this rapidly changing subject, including new chapters on turbulence and waves in the fluid outer core.

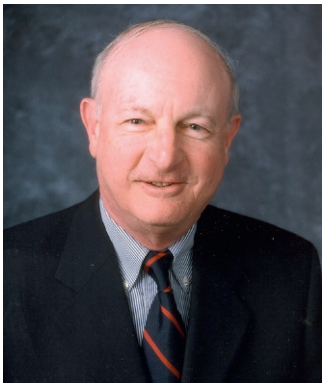
The volume on the Evolution of the Earth presents a new assessment on the role of volatiles in Earth evolution.

There are new chapters in the Physics of Terrestrial Planets and Moons volume on planetary atmospheres, planetary tectonics, extrasolar planets and asteroids and comets. The study of extrasolar planets has had explosive growth in the last decade and new spacecraft missions have significantly enhanced our knowledge of solar system planets and moons.

Finally, this is an opportunity for me to thank and acknowledge the editors of the individual volumes for their hard work, patience, dedication, and commitment in bringing this new edition of the Treatise to fruition. I am also indebted to Gemma Tomalin of Elsevier for her care and attention to the organization and production of the Treatise and to Deepa Arumugam for her careful coordination of the proofs from the copyeditor and typesetter. Of course, the authors of the individual chapters are the ones who are responsible for the excellence of the Treatise. We have all had some ups and downs in the process of producing the second edition of the Treatise but the final product is one that we can all view with pride.

Gerald Schubert
Editor-in-Chief

EDITOR-IN-CHIEF



Gerald Schubert is Distinguished Professor Emeritus in the Department of Earth, Planetary and Space Sciences at the University of California, Los Angeles. His research interests encompass the physics of the interiors and atmospheres of the Earth, the Moon, and the other moons and planets of the solar system. He is co-author with Donald Turcotte of *Geodynamics* (ed. 3, Cambridge University Press, 2014) and with Turcotte and Peter Olson of *Mantle Convection in the Earth and Planets* (Cambridge University Press, 2001). He is the author of over 540 research papers. He has participated in a number of NASA's planetary missions, including Apollo, Pioneer Venus, Magellan, and Galileo, and has been editor and editorial board member of many journals, including *Icarus*, *Journal of Geophysical Research*, *Geophysical Research Letters*, and *Annual Reviews of Earth and Planetary Sciences*. Professor Schubert is a Fellow of the American Geophysical Union and a recipient of the Union's James B. MacElwane medal and the Harry H. Hess medal. He is a member of the National Academy of Sciences and the American Academy of Arts and Sciences.

This page intentionally left blank

CONTRIBUTORS

E Barth-Naftilan
Yale School of Forestry and Environmental Studies, New Haven, CT, USA

Y Bernabé
Massachusetts Institute of Technology, Cambridge, MA, USA

ME Best
Bemex Consulting International, Victoria, BC, Canada

A Binley
Lancaster University, Lancaster, UK

V Chandler
Minnesota Geological Survey, St. Paul, MN, USA

DV Fitterman
U.S. Geological Survey, Long Island City, NY, USA

KL Ford
Geologist (Retired Physical Scientist, Geological Survey of Canada)

DR Glaser
Washington River Protection Solutions, LLC, Richland, WA, USA

PWJ Glover
University of Leeds, Leeds, UK

M Jackson
University of Minnesota, Minneapolis, MN, USA

MD Jackson
Imperial College, London, UK

TC Johnson
Pacific Northwest National Laboratory, Richland, WA, USA

PG Killeen
Geophysical Consultant (Retired Research Scientist, Geological Survey of Canada)

E Kragh
Schlumberger Gould Research, Cambridge, UK

R Laws
Schlumberger Gould Research, Cambridge, UK

Y Li
Colorado School of Mines, Golden, CO, USA

A Maineult
Institut de Physique du Globe de Paris, Paris, France

BM Moskowitz
University of Minnesota, Minneapolis, MN, USA

M Müller-Petke
Leibniz Institute for Applied Geophysics, Hannover, Germany

CJ Mwenifumbo
Consulting Geophysicist (Retired Research Scientist, Geological Survey of Canada)

M Nabighian
Colorado School of Mines, Golden, CO, USA

JD Phillips
U.S. Geological Survey, Denver, CO, USA

JOA Robertsson
ETH-Zurich, Zürich, Switzerland

DF Rucker
hydroGEOPHYSICS, Inc., Tucson, AZ, USA

JE Saiers
Yale School of Forestry and Environmental Studies, New Haven, CT, USA

DR Schmitt
University of Alberta, Edmonton, AB, Canada

L Slater
Rutgers University, Newark, NJ, USA

J van der Kruk
Agrosphere, Institute of Bio and Geosciences, Jülich, Germany

U Yaramanci
Leibniz Institute for Applied Geophysics, Hannover, Germany; Berlin Technical University, Berlin, Germany

11.01 Resources in the Near-Surface Earth: An Introduction and Overview

L Slater, Rutgers University, Newark, NJ, USA

© 2015 Elsevier B.V. All rights reserved.

This chapter is a revision of the previous edition chapter by W R Buck, Volume 6, pp. 335–376, © 2007, Elsevier B.V.

11.01.1	Introduction: Definitions, Opportunities, and Challenges	1
11.01.2	Physical Properties of the Near Surface	3
11.01.3	Hydrocarbon and Mineral Resources in the Near Surface	3
11.01.4	Anthropogenic Footprints in the Near Surface	4
11.01.5	The Cryosphere	4
11.01.6	Water Resources and Hydrogeophysics	5
11.01.7	Biogeophysics	8
11.01.8	Interpretation of Near-Surface Geophysical Datasets	9
11.01.9	Overview of the New Volume	10
11.01.10	Future Challenges for NSG	11
11.01.10.1	Unconventional Hydrocarbons	11
11.01.10.2	Environmental Stewardship	12
11.01.10.3	Carbon and Climate Change	13
11.01.11	Future Needs in NSG	14
References		17

11.01.1 Introduction: Definitions, Opportunities, and Challenges

Near-surface geophysics (NSG) was not specifically addressed in the first edition of the Treatise on Geophysics. The decision to include a new volume in the second edition of the Treatise on Geophysics on the near surface reflects the ever-growing contribution that geophysics is making to our understanding of the outermost shell of the Earth. This growth is evidenced by the formation of a new focus group in NSG within the American Geophysical Union, the largest Earth science organization worldwide, in 2006 (**Figure 1**).

The near surface is perhaps a somewhat nebulous term that means different things to different people. From a purely material properties perspective, one defining characteristic of the near surface is the presence of porosity and the mobile fluids filling it. These fluids increase the number of material properties needed to describe the geophysical (e.g., seismic and electrical) properties of Earth materials (e.g., see the discussion of seismic properties given by Schmitt, **Chapter 11.03**). However, the near surface can also be defined from a process-based perspective. For example, a critical defining metric for the near surface is the upper part of the Earth where humans interact with natural Earth systems. These interactions include the exploitation of valuable resources in the Earth and degradation of natural systems by pollution that inevitably accompanies industrialization and economic development. New interactions driven by the human response to the threat of climate change include the sequestration of carbon in subsurface reservoirs. In all cases, such human interaction modifies natural biogeochemical cycles and the physicochemical properties of Earth materials. Given that the world population is expected to reach more than nine billion by 2050 and with the rapid economic development and rising living standards in emerging economies, this human modification of natural systems can be expected to intensify. Understanding the effects of such human impacts is challenging as the near surface is a dynamic, nonlinear system characterized by multiple feedback mechanisms; consequently, our understanding of the long-term evolution of coupled human–natural systems is far from complete.

An important component of the near surface as defined in the preceding text is the 'critical zone' (CZ) (Brantley et al., 2006, 2007; Chorover et al., 2011), being the near-surface environment where water, rock, air, and life meet in a dynamic interplay that generates soils, sustains ecosystems, and shapes landscapes (NRC, 2001). The CZ is often assumed to extend from the uppermost periphery of vegetation to the lowest limits of freely circulating groundwater (**Figure 2**). Studies of the CZ are amplifying worldwide, but the majority of this research has focused on just the upper 1–2 m or so of weathered rock and soil in landscapes (Dietrich, 2010), even though it is well recognized that the top 1–2 m is often profoundly influenced by processes that occur beneath it to hundreds of meters or more in some landscapes (Chorover et al., 2007; Detheir and Lazarus, 2006; Graham et al., 2010). The 'deep CZ,' below the upper 1–2 m, has been referred to as the 'unmeasured zone' (Dietrich, 2010) because it is beyond the limits of soil augers and test pits. Yet, the deep CZ is actually highly accessible with well-established geophysical imaging technologies that are the topic of this volume. Such misrepresentation of the deep CZ highlights a need for better collaboration between near-surface geophysicists and the hydrologists, geochemists, and geomorphologists currently studying CZ processes and the near surface in general.

This volume is primarily focused on the geophysical principles and associated geophysical methods that are increasingly used to probe the near-surface Earth. As with solid Earth studies, NSG is motivated by the need for (usually noninvasive) data that can provide information to help improve our understanding of the functioning of the Earth. In addition, the development of NSG has been driven by the need to use geophysical

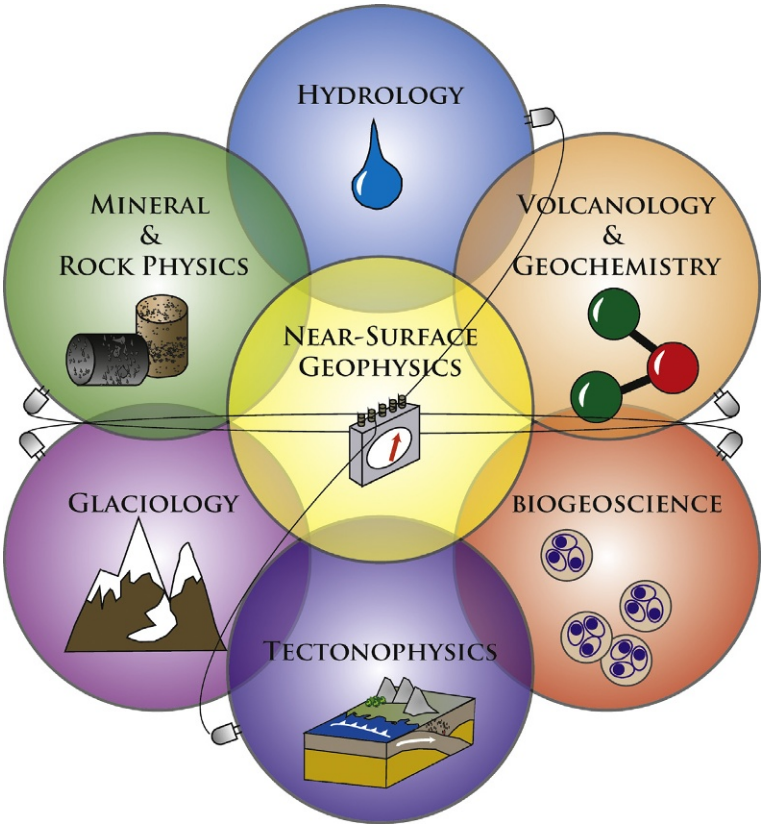


Figure 1 The concept of near-surface geophysics and its relation to other sections and focus groups of the American Geophysical Union (AGU). Reproduced from Slater L, Knight R, Singha K, Binley A, and Atekwana E (2006) Near-surface geophysics: A new focus group. *Eos*, *Transactions American Geophysical Union* 87(25): 249.

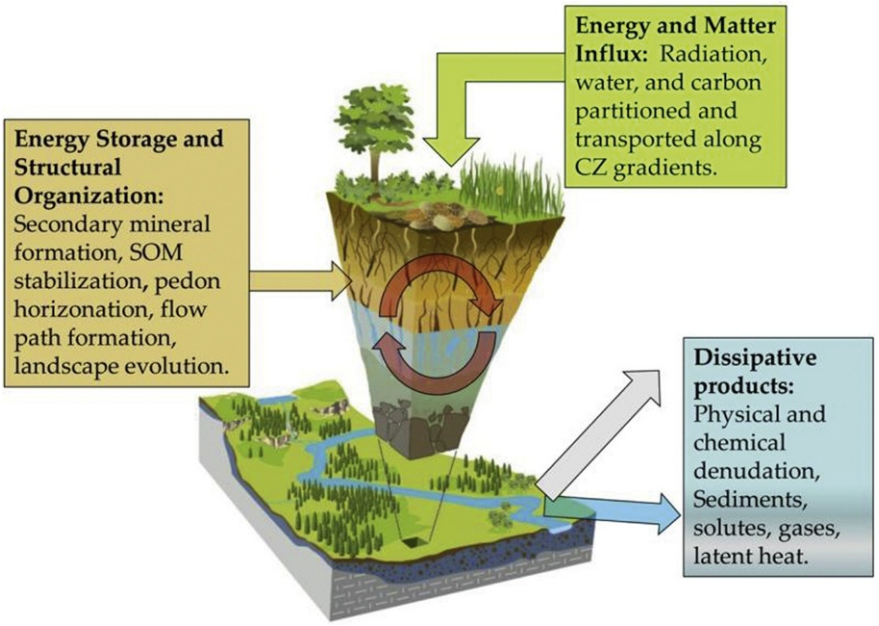

Energy and Matter Influx: Radiation, water, and carbon partitioned and transported along CZ gradients.
Energy Storage and Structural Organization: Secondary mineral formation, SOM stabilization, pedon horizonation, flow path formation, landscape evolution.
Dissipative products: Physical and chemical denudation, Sediments, solutes, gases, latent heat.

Figure 2 The critical zone of the near-surface Earth emphasizing the interaction of water rock and organisms. Modified from Chorover J, Kretzschmar R, Garcia-Pichel F, and Sparks DL (2007) Soil biogeochemical processes within the critical zone. *Elements* 3(5), 321–326; Chorover J, Troch PA, Rasmussen C, et al. (2011) How water, carbon, and energy drive landscape evolution and surface water dynamics: The Jemez River Basin – Santa Catalina Mountains Critical Zone Observatory. *Vadose Zone Journal* 10: 884–889 (special issue on the Critical Zone).

technologies to locate and exploit practically extractable resources that drive economic development. Although the underlying physical concepts employed to study the near surface are similar to solid Earth studies, the low pressure and temperature regime of the near surface means that the physical properties are distinctly different from the deep Earth. Similarly, the methods developed to investigate spatiotemporal variations in the properties of the near-surface Earth have distinct characteristics that have developed in response to the information needs defined over the last century. This volume therefore includes a comprehensive set of chapters dedicated to both the physical properties of the near-surface Earth and the state-of-the-art techniques that are used to provide information on this increasingly important subsystem of our planet. The volume also includes chapters describing examples of how geophysics provides information on extractable resources and environmental degradation resulting from human activities. The topical issues of water resources and unconventional fossil fuels are also discussed as they are emerging areas where NSG is primed to play an important future role.

11.01.2 Physical Properties of the Near Surface

The most important physical properties of the near surface required to understand the impact of human interactions on the Earth, and on CZ processes in particular, are characteristically distinct from the deep solid Earth. This difference is primarily driven by the presence of fluids that fill interconnected pores that are present in near-surface rocks and soils as a result of the low confining pressures. This open network of fluid-filled pores results in the near surface being highly porous and permeable to fluid flow. Permeability, which defines the ease with which a fluid can pass through this pore network, is a critically important transport characteristic in the near surface as it determines whether resources are economically viable to exploit and also exerts a primary control on the fate of contaminants in the subsurface. Bernabé and Maineult (Chapter 11.02) address the formidable rock physics task of modeling and predicting the permeability of near-surface porous media. The most porous and permeable rocks and sediments of the near surface form the vast aquifers and reservoirs that contain water and petroleum resources. The exploitation of such resources, along with understanding the impact of human degradation of these resources, requires a profound understanding of the physics controlling fluid flow in porous media and the petrophysical relations that link geophysical properties to pore space properties.

The presence of a pore network profoundly changes the electrical and seismic properties of the near surface relative to the deep Earth. Glover (Chapter 11.04) explains how the electrical properties of the near surface are primarily determined by electrolytic conduction occurring through pores, and in contrast to the deep Earth, the minerals (excluding metallic mineral deposits) are poor conductors and considered insulators. Schmitt (Chapter 11.03) similarly describes how the seismic properties of near-surface materials are fundamentally controlled by the pore space as a result of the strong difference in seismic velocities of pore-filling fluids and the mineral matrix. The magnetic properties of the near-surface Earth are also distinctly different from those of the deep Earth since they are affected by erosion and deposition and multiple biogeochemical processes controlling iron cycling. In addition, various anthropogenic processes alter the magnetic properties of near-surface materials. An exhaustive review of the magnetic properties of near-surface Earth materials is presented by Moskowitz (Chapter 11.05).

11.01.3 Hydrocarbon and Mineral Resources in the Near Surface

The major petroleum resources of the Earth are located in unique geologic structures (sedimentary basins) that can be imaged with high-resolution seismic methods (see Robertsson et al., Chapter 11.06). The petrophysical relationships that link geophysical properties to pore space properties have allowed geophysicists to evaluate the economic value of petroleum resources. Glover (Chapter 11.04) highlights the example of Archie's law, which provides an interpretation of the electrical properties of petroleum reservoirs in terms of the interconnected porosity controlling the economic viability of extracting the resource. According to Glover, it is thanks to Archie's law that "over 1.5 trillion barrels of oil reserves and 7.5 trillion standard cubic feet of gas have been quantified and produced, running the global economy for the last 70 years."

The major mineral resources of the Earth also have distinct geophysical properties (elastic, electromagnetic (EM), and gravitational) that are related to the physicochemical properties defining the quality (e.g., energy or precious metal content) of these resources (see Best, Chapter 11.15). The widescale adoption of geophysical exploration methods for locating major resources has therefore played a critical role in the economical and industrial development of nations. Perhaps the most striking example is the almost complete reliance of the oil exploration industry on marine seismic methods for locating potential petroleum resources worldwide. As reviewed by Roberttsson et al. (Chapter 11.06), the complexity of these methods has grown dramatically in the last few decades in response to the urgent need to continue to locate previously untapped petroleum reservoirs and unconventional reserves. Similarly, geophysics has often provided the information needed to delineate the major mineral resources of the Earth. This volume includes two lengthy chapters describing the importance of geophysics in the location of key mineral resources. Whereas Killeen et al. (Chapter 11.14) provide an in-depth description of the contribution of radiometric methods to mineral exploration, Best (Chapter 11.15) provides a comprehensive overview of examples of the use of a broad range of geophysical techniques for locating a wide range of mineral resources.

Human interaction with the near-surface Earth results in changes to geophysical properties through resource exploitation, economic development, and environmental pollution. The geophysical properties of the near-surface Earth are therefore highly dynamic and subject to changes over short timescales relative to the deeper Earth where geophysical properties are approximately constant over much larger timescales. This human alteration of the geophysical properties of the near

surface provides a signature that can be detected and monitored with geophysical methods.

Since the end of the Second World War, geophysical instrumentation and sensors have rapidly evolved to permit noninvasive or minimally invasive characterization and monitoring of the near-surface Earth at unprecedented spatiotemporal scales. Evidence of this evolution is most dramatic in the petroleum exploration industry, where development of the marine seismic method has been fueled by the remarkable successes of this technique in locating major petroleum reservoirs worldwide (Robertsson et al., Chapter 11.06). In the field of mineral exploration, airborne geophysical methods (particularly EM and magnetics) have been extensively developed to map mineral resources over large spatial areas (Best, Chapter 11.15).

11.01.4 Anthropogenic Footprints in the Near Surface

At a somewhat different scale, the same physics has been employed to target a legacy of a very different human interaction with the subsurface: the plethora of unexploded ordnance (UXO) that litters the very near surface of countries that is the footprint on the near surface of modern human conflict. Custom-built EM and magnetic sensors have been developed to have optimal sensitivity to shallow targets with strong magnetic susceptibility and electrical conductivity contrasts with surrounding soil, representing different types of UXO (shape and physical property contrast) (e.g., Huang and Won, 2003). Specific algorithms have been developed to help discriminate likely UXO targets from non-UXO targets (Pasion and Oldenburg, 2001). This application of geophysics remains an area of active research as not all UXO is alike from a geophysical perspective; for example, whereas metallic landmines are relatively easy to detect, small plastic explosives remain a major challenge for reliable geophysical detection. The biggest challenge is being able to discriminate UXO from other inert metallic debris that also litters the near surface of conflict zones.

Archaeology is yet another near-surface scientific discipline where geophysics has facilitated important advances. For decades, archaeologists have been adopting geophysical methods to assist in the delineation of the footprint in the near surface of former civilizations. Geophysical methods have repeatedly proved invaluable for delineating the distribution of buried artifacts prior to the commencement of an excavation (Scollar et al., 1990). In some instances, geophysical imaging of archaeological sites has been performed to illuminate the past without the need for excavation, which is often destructive. A wide range of geophysical techniques covered in this volume have been routinely used for archaeological prospecting, including resistivity, EM, magnetometry, ground-penetrating radar (GPR), and seismics. Archaeological prospecting with geophysics may involve relatively simple data acquisition and minimal processing that can be performed by a nonexpert geophysicist. For example, a simple four-electrode tool has been developed for reconnaissance surveys, although resistivity imaging methods are becoming more popular (Papadopoulos et al., 2006). Figure 3 is a visually appealing example of the rich archaeological information obtainable from high-resolution magnetometry (Weymouth, 1986). This relatively simple magnetic survey was performed over Big Hidatsa Village, an Indian settlement that was occupied from the late 1700s to the early 1800s and located in the Knife River Indian Villages National Historic Site in North Dakota. The resulting magnetic anomaly map reveals the distribution of house sites, marked by circular anomalies, several of which are not visible from surface topography. The magnetic anomaly at the center of most house sites is associated with thermoremanent magnetism of a central fire pit (see Moskowitz, Chapter 11.05). The high magnetic regions between houses are attributed to trash middens, where the high organic content in the middens is caused by the reduction of hematite to magnetite (see Moskowitz, Chapter 11.05). Other anomalies on some house floors were found to be associated with storage pits, pieces of iron, or heated boulders as confirmed from test excavations. In the top right of the image, two parallel lines are interpreted as the magnetic response of wagon ruts. Overall, a simple geophysical survey has provided a remarkably detailed reconstruction of this relict village.

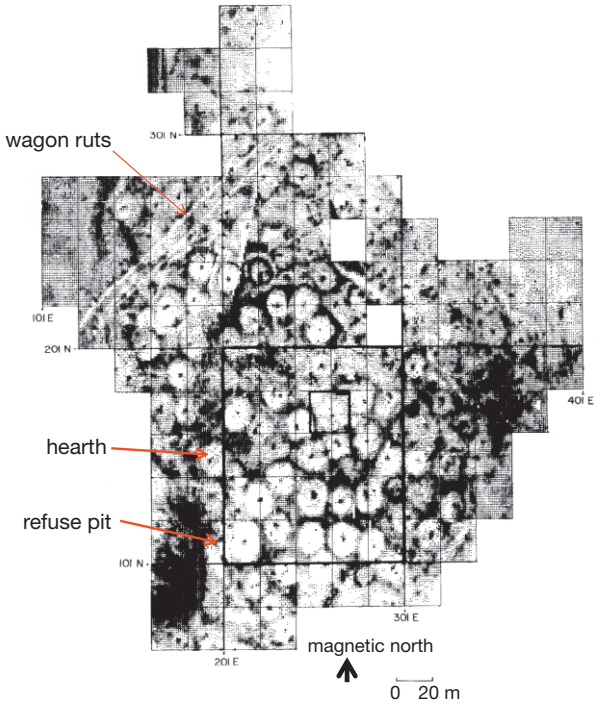


Figure 3 Result of a magnetic survey over Big Hidatsa Village, an Indian village that was occupied in the late 1700s to the early 1800s and located in the Knife River Indian Villages National Historic Site in North Dakota. The magnetic survey reveals the configuration of the houses making up the village. The image is made up of ten symbols spanning a range of magnetic anomaly from −3 to 15 nT, with increasing size/darkness representing an increasing anomaly. Reproduced from Weymouth JW (1986) Archaeological site surveying program at the University of Nebraska. *Geophysics* 51(3): 538–552.

11.01.5 The Cryosphere

Yet another area of active near-surface geophysical research relates to the study of the cryosphere, defined as the portions

of the Earth where water is in the solid form. The challenges of drilling through ice, permafrost, or temporarily frozen ground are substantial; consequently, cryosphere researchers are increasingly turning to geophysical imaging in order to obtain subsurface information on this unique subsystem of the Earth that is particularly sensitive to climate change. Near-surface geophysical data have been extensively used to image the thickness of ice sheets; to obtain information on the morphology, physical properties, and stratigraphy of deposits below ice sheets; and to obtain information on their internal structure. As one of many examples, seismic data over an ice stream of the marine ice sheet of West Antarctica identified a saturated layer beneath this ice stream that led to new understanding of the physical mechanisms governing the movement of ice streams (Blankenship et al., 1986). Further seismic work in West Antarctica showed compelling evidence of the role of subglacial geology in controlling the position of ice streams (Anandakrishnan et al., 1998).

Seismic and GPR data have similarly been effectively employed to image subglacial and englacial features that control the dynamics of glaciers. GPR is a particularly effective method for imaging glaciers due to the low attenuation of the EM waves in ice and the strong contrasts in dielectric permittivity that arise due to heterogeneities within and below glaciers. Glacier beds are readily detected with low-frequency (10–50 MHz) GPR, and repeated measurements have been used to monitor the shrinkage of temperate glaciers (Ramirez et al., 2001). The internal structure of rock glaciers has also been extensively studied with GPR and used to infer structural development of these landforms (Berthling et al., 2000). Surface and cross-borehole GPR measurements have been extensively used to understand englacial hydrology, for example, the distribution of substantial volumes of water in glacier ice (Murray et al., 2000).

Subglacial drainage conditions have been noninvasively assessed using self-potential measurements from which estimates of subglacial flow velocity and direction have been derived (Kulessa et al., 2003). Helicopter-borne GPR has been used to determine spatial variability in snow accumulation within Alpine glaciers over large areas (Machguth et al., 2006). The acquisition of active source shallow seismic data has also been effective in imaging the internal structure and bed properties of Alpine glaciers (Musil et al., 2002). Passive seismic methods have been used to detect and monitor rapid retreat of glaciers associated with sudden calving events (O'Neel et al., 2007). Qamar (1988) showed evidence for a simple relationship between the volume of an iceberg breaking off the glacier terminus and the duration of the seismic signal generated, using it to estimate the calving rate of the Columbia Glacier (Alaska) from seismological data alone.

More recently, NSG has contributed to understanding the dynamics of permafrost thawing, which could have significant implications for global climate change due to the large quantities of methane locked up in organic-rich soils in permafrost (Christensen et al., 2004; Walter et al., 2006). GPR (Hinkel et al., 2001) and DC resistivity tomography (Hauck and Mühll, 2003) both have been extensively utilized to characterize spatial variability in subsurface permafrost. Thermokarst lakes, being lakes filling depressions caused by continuous sediment subsidence due to thawing ground, are hotspots of methane emissions (Walter et al., 2006) and thus of great interest with respect to predicting the changing carbon dynamics of permafrost regions in response to climate warming. GPR and shallow seismic profiling have resolved permafrost structures below these thermokarst lakes (Schwamborn et al., 2002a,b), providing information on the geometry of the talik (thaw bulb) associated with these lakes. **Figure 4** shows the results of seismic reflection and electrical resistivity imaging datasets that provide unique information on the geometry of the talik underlying two thermokarst lakes. Most recently, Parsekian et al. (2013) used nuclear magnetic resonance (NMR) soundings to detect depth of the talik and to infer connectivity of the talik to the regional groundwater (see Figure 20 of Müller-Petke and Yaramanci, **Chapter 11.13**). Thaw bulb geometry depends on lake age and also provides information on whether the thaw bulb is at steady state, expanding (thawing), or shrinking (West and Plug, 2008). Consequently, geophysics has the potential to provide critical data required to understand mechanisms and rates of thaw bulb expansion.

Recently, airborne EM imaging has been used to acquire unique information on the large-scale variability in subsurface permafrost in the Yukon Flats, Alaska (Minsley et al., 2012). The results of this multifrequency domain EM survey along an approximately 1800 km line of helicopter-borne data yielded remarkable information on the spatial configuration of present-day permafrost and also provided insights into the interactions of groundwater with permafrost. An example of the rich information content in this dataset is provided in **Figure 5**, which shows three-dimensional cutouts of this massive EM dataset around two locations of significant interest with respect to interactions of permafrost and hydrogeology. The first (**Figure 5(a)**) is around Twelve Mile Lake where the geometry of the thaw bulb below the lake is again of interest. The second (**Figure 5(b)**) is in the vicinity of the Yukon River, where the distribution of river talik and the lateral migration of this river are of interest. The authors emphasized that this spatially rich dataset represents a critical baseline against which future repeat EM surveys over the same area will be able to provide unique large-scale information on the change in the configuration of permafrost in response to climate change.

11.01.6 Water Resources and Hydrogeophysics

Over the last two decades, there has been a dramatic growth in the development of geophysical technologies and sensors primarily for the characterization of water resources and for monitoring the impacts of pollution on the subsurface Earth. The increasing challenge of quantifying extractable, economically viable, potable water supplies is a major factor driving this growth and has led to the definition of a new subdiscipline of hydrology known as *hydrogeophysics* (Rubin and Hubbard, 2005). According to the American Geophysical Union, "Hydrogeophysics involves use of geophysical measurements for estimating parameters and monitoring processes that are important to hydrological studies, such as those associated with water resources, contaminant transport, ecological and climate investigations." Freshwater resources are fundamental for maintaining human health, agricultural production,

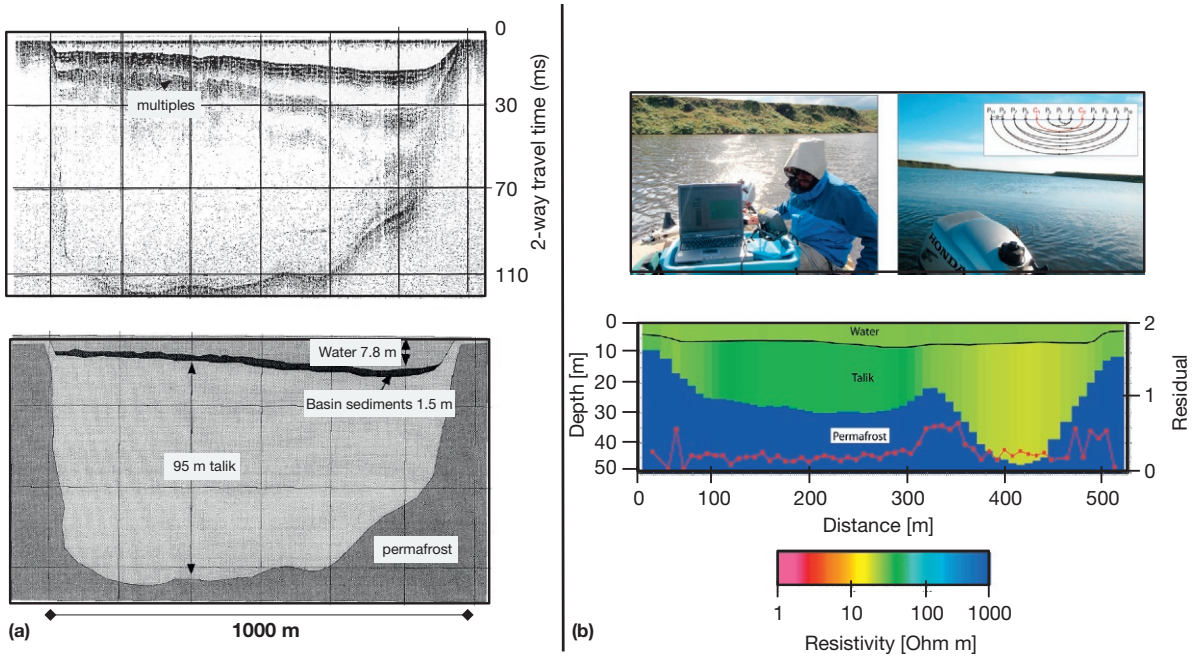


Figure 4 Two examples of geophysical imaging of thaw bulb dimensions under thermokarst lakes: (a) shallow seismic reflection data (top) acquired over Lake Nikolay, Arga Island, western Lena River delta, and resulting interpretation of bulb dimensions (bottom); (b) continuous waterborne resistivity data acquisition over Lake Claudi, Seward Peninsula (top), and talik thickness inferred from a 1-D laterally constrained inversion (bottom). Red line in bottom figure is data fitting residual. (a) Modified from Schwamborn G, Andreev A, Rachold V, et al. (2002) Evolution of Lake Nikolay, Arga Island, Western Lena River delta, during Late Pleistocene and Holocene time. *Polarforschung* 70: 69–82; (b) unpublished data.

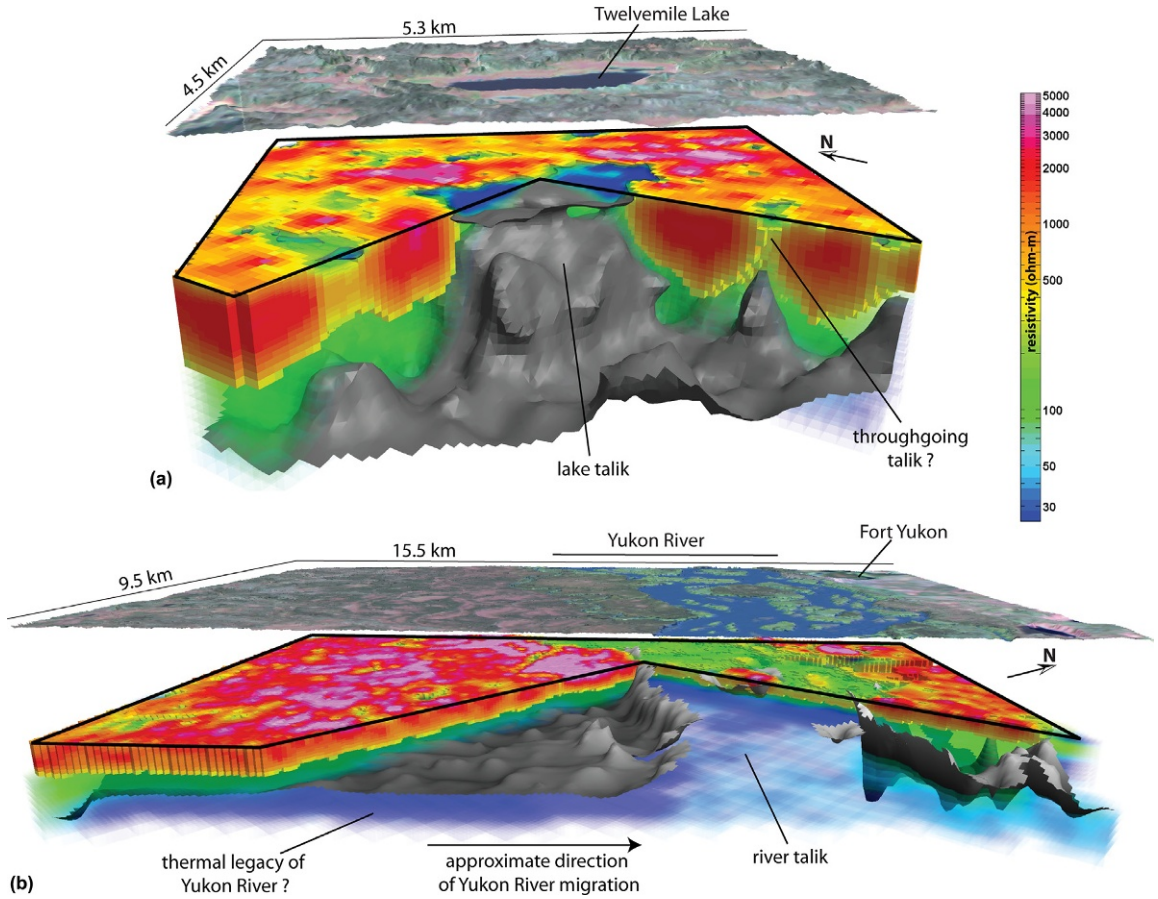


Figure 5 Examples of the unique information content obtained by Minsley et al. (2012) from a multifrequency airborne EM survey covering 16 line km of the Yukon Flats near Fort Yukon, Alaska: (a) survey cutout in the vicinity of Twelve Mile Lake highlighting inferred permafrost distribution (gray) and geometry of lake talik; (b) survey cutout in the vicinity of the Yukon River showing distribution of river talik and interpreted direction of Yukon River migration. Reproduced from Minsley BJ, Abraham JD, Smith BD, et al. (2012), Airborne electromagnetic imaging of discontinuous permafrost. *Geophysical Research Letters* 39: L02503.

economic activity, and critical ecosystem functioning. In this respect, freshwater is the most valuable resource on Earth since, unlike in the case of fossil fuels, no alternatives exist. The concept of peak oil (a peaking and subsequent decline in oil production) has recently been adopted to evaluate vulnerability of water resources. Gleick and Palaniappan (2010) argued that peak nonrenewable water occurs in groundwater systems when withdrawal rates substantially exceed natural recharge rates and where overpumping or contamination leads to a peak in production followed by a decline, similar to peak-oil curves. It is widely recognized that global human appropriation of freshwater could outpace freshwater availability (Postel et al., 1996). The increasing concern over the availability of freshwater, elevated by the threat to water resource availability associated with climate change coupled to population growth, is a driving force for the development of imaging techniques that can locate (in developing countries) and monitor (in developed countries) potable aquifers.

Geophysics is consequently increasingly making contributions to improving understanding of subsurface hydrogeology, including groundwater flow and solute transport, moisture dynamics in the unsaturated zone, groundwater–surface water interaction (and its effect on ecological systems), and freshwater–saltwater exchange. Land-based electrical and EM methods have been advanced to permit high-resolution 3-D (and 4-D in time) imaging of aquifers and the transport of solutes and contaminants. As an example of the hydrogeologic information content of geophysical datasets, Figure 6 shows the evolution of a mildly saline tracer through the unsaturated zone at a sandstone site imaged by Winship et al. (2006) due to the decreases in resistivity resulting from the passage of more conductive fluid through the interconnected pore space (see Glover, Chapter 11.04, for a review of near-surface electrical properties). In this example, borehole electrical tomographic imaging methods based on the measurement of a large number of voltage differences (ΔV) associated with the electric field due to an injected current (I) (Figure 6(a)) provide information on aquifer recharge rates and solute transport through the vadose zone (Figure 6(b)). The tracer migrates relatively rapidly through the unsaturated profile until further vertical flow is impeded by a low-permeability layer. Although changes in electrical properties offer many opportunities to study hydrogeologic processes, other geophysical properties are increasingly being monitored for hydrogeologic information. For example, high-resolution time-lapse gravity methods have been recently employed to monitor changes in the volume of water stored in aquifers or to monitor aquifer storage and recovery projects (Davis and Batzle, 2008).

Hydrogeophysical data are increasingly being acquired at watershed and larger scales as a result of recent advances in airborne geophysical technologies. Airborne geophysics has been fundamental for locating and characterizing mineral resources and a comprehensive set of mineral exploration case studies is presented by Best (Chapter 11.15). Geophysical techniques that are particularly well suited to mapping 2-D variations in geophysical properties from aircraft over large scales include magnetics, radiometrics (see the extensive review by Killeen et al., Chapter 11.14), and EM as described by

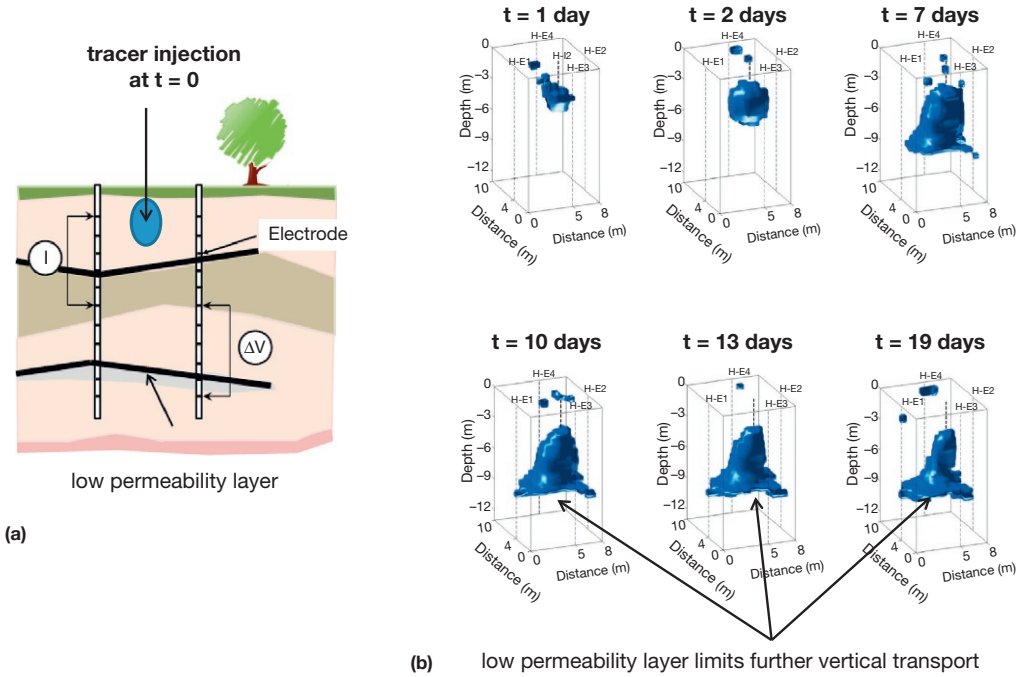


Figure 6 Example use of electrical geophysics to resolve transport characteristics of near-surface unsaturated sandstone: (a) Cross-borehole electrical imaging whereby voltage differences (ΔV) are recorded in response to current (I) injection is performed during the application of a mildly saline tracer; (b) 3-D tomographic images of changes in resistivity due to an increase in salinity and moisture capture the transport properties of the site, including the effect of a low-permeability layer that impedes further vertical transport. Modified from Winship P, Binley A, and Gomez D (2006) Flow and transport in the unsaturated Sherwood sandstone: Characterization using cross-borehole geophysical methods. *Geological Society, London, Special Publications* 263(1): 219–231.

Fitterman (Chapter 11.10). A very recent development is airborne time-domain EM (TEM) that provides an opportunity to obtain 3-D variations in geophysical properties over large spatial scales. A striking example of the hydrogeophysical information content of such airborne TEM data is given in d'Ozouville et al. (2008), where TEM was used to map the hydrogeologic structure of the supposedly 'waterless' Santa Cruz Island in the Galápagos Islands (Figure 7). The survey provided unprecedented resolution of the 3-D structure of the island, revealing the location of previously unknown perched aquifers and the geometry of the saltwater–freshwater interface within the basal aquifer of the island. The ability to capture such rich, high-resolution hydrogeologic information across such large spatial scales has wide-reaching implications for improving the understanding of a range of processes occurring in the near-surface Earth.

An increasingly important subdiscipline within hydrogeophysics is agricultural geophysics, wherein near-surface techniques with high resolution are increasingly being employed to evaluate soil moisture state and irrigation practices (Allred et al., 2008). In order to feed the projected 9 billion people by 2050, the Food and Agriculture Organization (FAO) of the United Nations (UN) projects a required increase of 65–70% in cereal grains (*Nations* 2009). Currently, 70% of global water consumptive use is for agriculture, with irrigation agriculture accounting for 40% of the global food production (Mekonnen and Hoekstra, 2011). Global trends in consumptive water use highlight a growing yet unsustainable reliance on groundwater resources (Mekonnen and Hoekstra, 2011). On the global scale, Clay (2004) estimated that 60% of the 2500 trillion liters used for agriculture each year are wasted through inadequate water conservation, losses in distribution, and inappropriate times and rates of irrigation.

In developed nations, geophysical technologies are increasingly being adopted in precision agriculture to assist with management of irrigation rates via zone management of sprinkler systems. Industry practice is increasingly including continuous mapping of soil apparent electrical conductivity (ECa) using commercially available technologies that can be pulled across fields. For example, Pan et al. (2013) used regression modeling with static ECa zones and elevation to help guide variable rate irrigation decisions throughout the season. Lunt et al. (2005) reported on the use of GPR reflection travel time data to estimate changes in soil water content under a range of soil saturation conditions throughout the growing season at a California winery. This work identified the potential of geophysics to support management decisions based on precision viticulture.

11.01.7 Biogeophysics

More recently, the term 'biogeophysics' has been used to describe an even newer Earth science subdiscipline concerned with the geophysical signatures of microbial interactions with geologic media (Slater and Atekwana, 2013). The majority of biogeophysics research to date has focused on the near-surface Earth given the need to monitor biogeochemical transformations associated with the natural attenuation and/or enhanced remediation of groundwater contamination. Biogeophysics

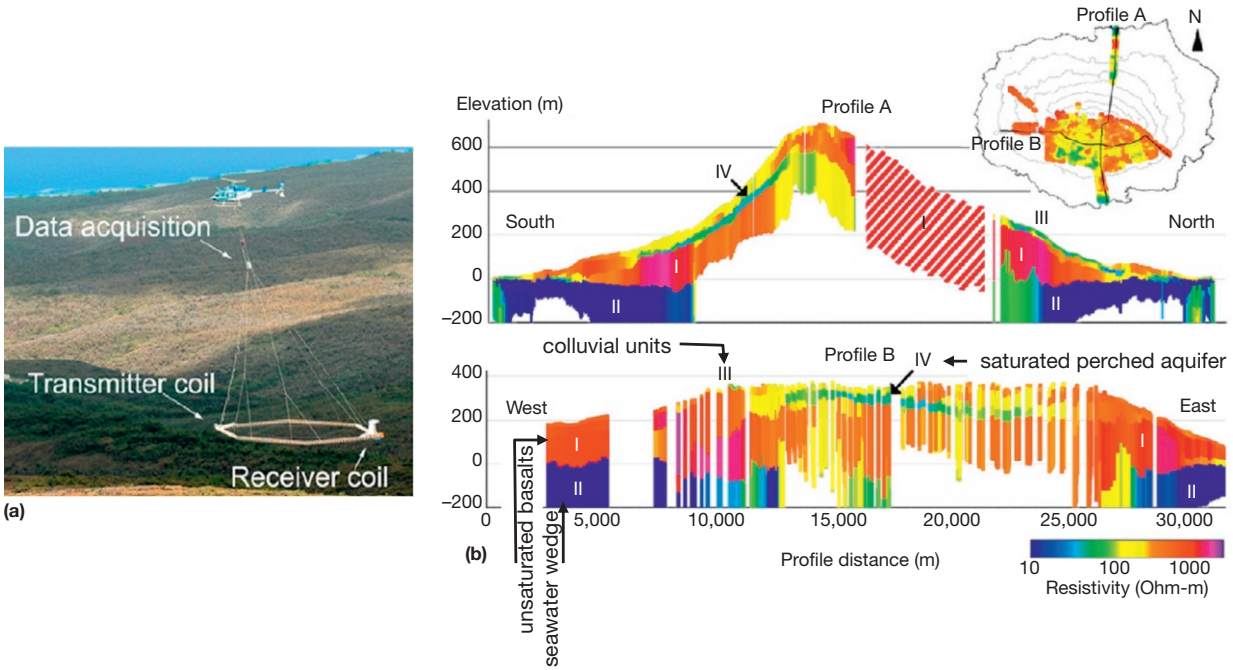


Figure 7 Application of airborne time-domain electromagnetics (TEM) to map the internal resistivity structure of Santa Cruz Island within the Galápagos Islands: (a) the SkyTEM system in operation showing TEM instrumentation hanging 30 m below a helicopter; (b) 2-D resistivity profiles across the island and interpretation of the associated hydrogeologic framework of the island, including perched aquifers and extent of saline intrusion in basal aquifer. Modified from d'Ozouville N, Auken E, Sorensen K, et al. (2008) Extensive perched aquifer and structural implications revealed by 3D resistivity mapping in a Galapagos volcano. *Earth and Planetary Science Letters* 269(3): 518–522.

research in the last decade has confirmed the potential for geophysical techniques not only to measure the physical and chemical properties of the near subsurface, as already well established, but also to detect microbes, microbial growth, and microbe–mineral interactions, thus representing a major paradigm shift in geophysical thinking (Atekwana and Slater, 2009). **Figure 8** illustrates how geophysical signatures can arise in response to such microbial processes, in this case the formation of biominerals on the sulfate-reducing bacteria *Desulfovibrio vulgaris*. In a two-step process involving the production of hydrogen sulfide and subsequent reactions with Fe^{2+} in solution, FeS is precipitated out on the surface of the cells, and the cells form FeS crusts on mineral grains (**Figure 8(a)**). These FeS crusts enhance a low-frequency electrochemical polarization phenomenon (**Figure 8(b)**) that is measured with the induced polarization method (see Binley, **Chapter 11.08**). **Figure 8(c)** shows the resulting changes in this electrical polarization due to the accumulation of such biominerals in laboratory columns as discussed in Slater et al. (2007).

The induced polarization geophysical method has recently been used to image and monitor this biogeochemical process in field experiments designed to track the effectiveness of such biomineralization for the remediation of radionuclide-contaminated sites (Flores Orozco et al., 2011). Additional examples of the geophysical signatures that can result from microbial processes are discussed by Glover (**Chapter 11.04**) and Jackson (**Chapter 11.09**). Considering hydrogeophysics and biogeophysics together, the monitoring of dynamic geophysical properties therefore provides unique opportunities to evaluate the human impact on near-surface systems.

11.01.8 Interpretation of Near-Surface Geophysical Datasets

One distinct difference between the deep (solid) Earth and the near-surface Earth is that, whereas geophysical data may be the only information obtainable from the deep Earth, geophysical data from the near-surface Earth must often be interpreted in the context of other existing datasets. An important area of research in NSG is coupled inversion (sometimes known as joint inversion) where multiple datasets (not necessarily geophysical) are integrated to reduce uncertainty and better constrain the estimation of subsurface properties or states compared to inversion conditioned to a single data type alone. In petroleum exploration, simultaneous joint inversion of EM, gravity,

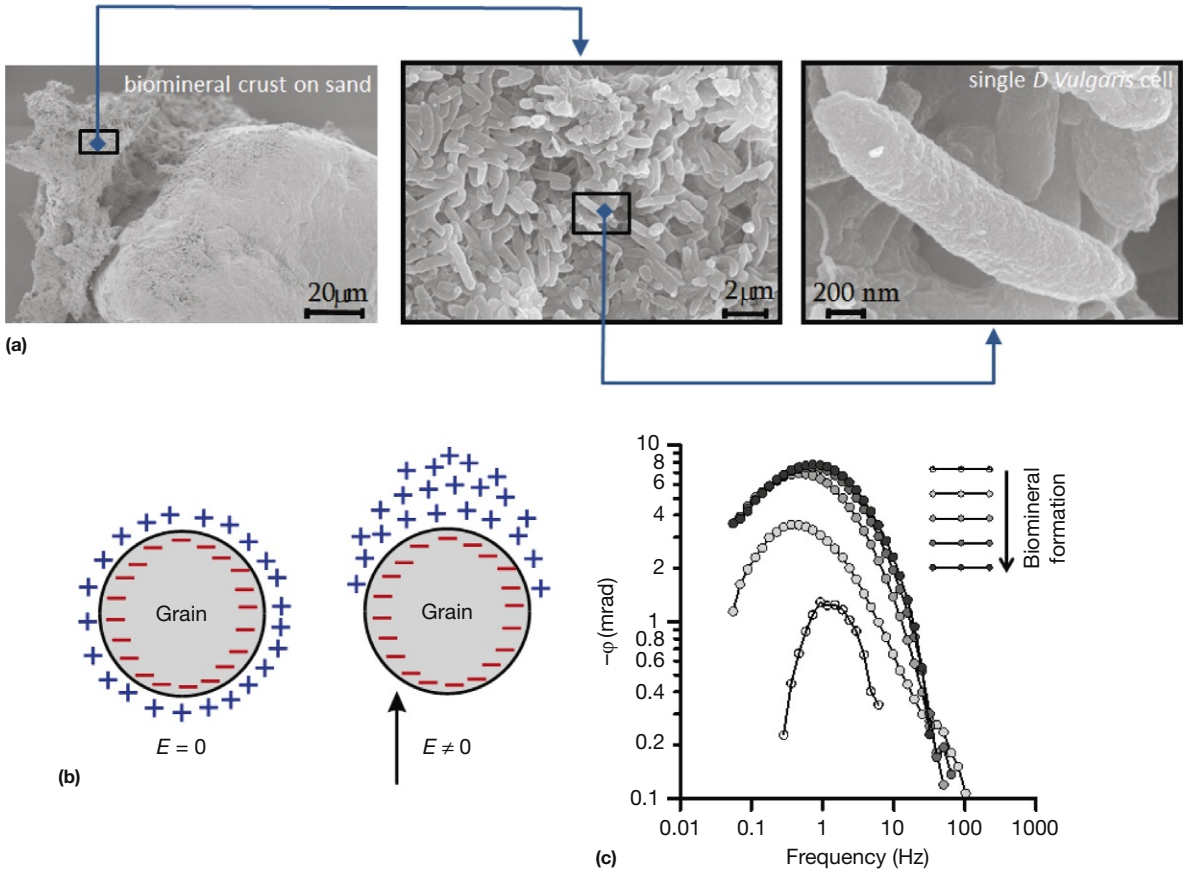


Figure 8 Geophysical signatures that can result from biomineralization processes in porous media: (a) Scanning electron microscope (SEM) images of *Desulfovibrio vulgaris* biomineralization at three scales; (b) the electrical polarization phenomenon enhanced by such biomineral formation (E = electric field); (c) the resulting frequency-dependent induced polarization geophysical signatures. Figures partly modified from Slater L, Ntarlagiannis D, Personna YR, and Hubbard S (2007) Pore-scale spectral induced polarization signatures associated with FeS biomineral transformations. *Geophysical Research Letters* 34(21): L21404.

and seismic data is increasingly conducted (Colombo et al., 2008). Simultaneous joint inversion has also gained considerable traction in the hydrogeophysics community, where geophysical and hydrologic datasets are coupled to improve the estimation of hydraulic properties, such as permeability. Other types of data fusion are also increasingly being employed, for example, the inclusion of hard geologic data from borehole logging to constrain geophysical inverse models. Van der Kruk (Chapter 11.07) describes how joint inversion strategies are enhancing the information content in GPR data.

Another distinct trend in NSG is toward 4-D data acquisition for temporal monitoring of subsurface transport and transformation processes over multiple spatial scales. This trend is apparent in multiple methods chapters submitted to this volume, whereby the historic development of the technique for generating 1-D or 2-D models of Earth structure is described, the focus then turning to 3-D or 4-D acquisition and processing. Methods and techniques originally developed for hydrocarbon exploration are now being employed in hydrogeophysics and environmental applications. However, whereas hydrocarbon exploration has encouraged the development of 4-D seismics to monitor changes in oil reservoirs (see Robertsson et al., Chapter 11.06), hydrogeophysical and environmental monitoring needs have largely encouraged the development of resistivity and EM methods for monitoring moisture dynamics and solute transport (Figure 6) (see chapter by Binley, Chapter 11.08, and Fitterman, Chapter 11.10). Even so, not all methods have matured to the point that 3-D and 4-D data acquisition is the norm. For example, the NMR method is an emerging geophysical technique (see Müller Pekte and Yaramanci, Chapter 11.13) that is currently largely limited to 1-D or 2-D data acquisition and processing.

Another technological development motivated by oil exploration and now making the transfer to hydrogeophysical and environmental applications is full waveform inversion where the full attributes of the propagated waveform are modeled instead of just first arrival times and amplitudes. Robertsson et al. (Chapter 11.06) explain that the introduction of full waveform inversion (Tarantola, 1984) provided the major breakthrough in imaging of complex oil reservoirs in the early to mid-2000s. As noted by van der Kruk (this volume), full waveform inversion currently represents the state-of-the-art interpretation of GPR data for hydrogeophysical and environmental investigations in complex media. Robertsson et al. (Chapter 11.06) provide striking examples of the additional information on subsurface structures that can be obtained from the full waveform inversion relative to conventional ray-based inversion of direct arrivals. An example of the improved resolving power of GPR when full waveform inversion is employed is presented in van der Kruk (Chapter 11.07).

11.01.9 Overview of the New Volume

This new volume of the Treatise on Geophysics was motivated by the increasing importance of the study of the field of geophysics for understanding the near-surface Earth and the omission of the near surface from the first edition of the Treatise on Geophysics. The primary objectives of this volume are (i) to define the geophysical properties of the near surface as they pertain to understanding human interactions with the Earth, (ii) to present the state-of-the-art geophysical exploration technologies as they pertain to characterizing and monitoring the near-surface Earth, and (iii) to relate geophysical properties and measurements to major applications areas where geophysical sensing has facilitated improved understanding. Table 1 summarizes this link between physical properties, measurement methods, and application area, where a color scale is used to describe relative importance in each application area (red = extensively used, white = minimal use).

The major sections of this volume on the near-surface Earth are summarized in Figure 9. The volume begins with a review of the physics of porous media as it pertains to near-surface soils and rocks. This chapter focuses in on the major challenge of modeling and predicting the permeability of Earth materials using the concepts and theory recently developed around length scales defining pore connectivity. The fundamental geophysical properties of the near-surface Earth are next introduced with dedicated chapters on electrical, seismic, and magnetic properties where the theoretical and empirical relations linking these geophysical properties to the physical and chemical characteristics of the near surface are described.

The second section of this volume contains comprehensive reviews on the exploration and monitoring methods that have evolved to obtain information on the spatial and temporal variations in the geophysical properties of the near surface. The volume includes chapters on the well-established geophysical investigation techniques, namely, seismic, galvanic electrical, EM, magnetic, and gravitational methods. These methods are also heavily utilized for investigating the structure of the deep Earth as discussed elsewhere in the Treatise. However, the spatial and temporal resolution demanded by near-surface investigations has required the development of measurement and processing methods that are uniquely different from those applied to the deep Earth. This section of the volume also contains dedicated chapters on geophysical methods that are unique in that they are almost exclusively applied to the near-surface Earth. These are the radiometric, self-potential, GPR, and NMR methods. Two of these techniques, GPR and NMR, were only developed in the 1970s and represent very recent additions to the suite of geophysical sensing technologies. Chapter 11.13 represents a review of the NMR method that is currently unparalleled elsewhere.

The volume next turns to address the increasing pressures resulting from human interactions with the Earth. Specifically, we consider some of the major economic and environmental challenges that require a better understanding of the geophysics of the near-surface Earth. Geophysical investigation of the near surface has been largely dominated by the need to locate conventional fossil fuel reserves and mineral deposits. Similarly, the location of new, less accessible fossil fuel reserves and mineral deposits requires a continued development of geophysical methods and improved understanding of how the physicochemical characteristics of rocks control geophysical properties. The extensive use of geophysics for mineral resource exploration, including recent developments, is covered in Chapter 11.15, with additional information found in Chapter 11.14.

Table 1 Relationship between physical properties, geophysical methods, and primary areas of application for studying the near-surface Earth

		APPLICATION					
PHYSICAL PROPERTY	**METHOD**	Petroleum exploration	Water resources	Minerals	Cryosphere	Environmental	Archaeology and UXO
Seismic velocity	*Marine and land seismics, passive seismic*						
Density	*Gravimetry*						
Electrical resistivity	*DC resistivity, Electromagnetics*						
Magnetic susceptibility, remanence	*Magnetometry*						
Dielectric permittivity	*Ground penetrating radar*						
Nuclear spin	*Nuclearmagnetic resonance (NMR)*						
Interfacial polarization	*Induced polarization*						
Radioactive decay	*Radiometrics*						

Color bar indicates relative importance of the technique to the application (red, high; white, low).

11.01.10 Future Challenges for NSG

11.01.10.1 Unconventional Hydrocarbons

The pressing need to locate dwindling conventional petroleum resources, combined with the necessity to identify and exploit unconventional fossil fuels, further demands an improved understanding of the geophysical properties of the near surface, along with new methodological developments. Unconventional hydrocarbons can be defined as "solid, liquid or gaseous hydrocarbons that have been generated over geological time from organic matter in source rocks, and which may also have been modified subsequently, especially in the case of solids and extra-heavy liquids" (Chew, 2014). Examples include bitumen (solid) in sandstone/carbonate, extra heavy oil (liquid), and tight gas in sandstone/limestone and gas in shale. They are trapped in an unconventional fashion and unconventional exploration and production methods are needed to locate them (Chew, 2014). Land-based exploration for unconventional fuels requires mapping of low-relief structures and stratigraphic traps and improved data quality for accurately placing horizontal wells (Keho and Kelamis, 2012). Seismic data acquisition and processing concepts developed for 3-D marine data acquisition (Robertsson et al., **Chapter 11.06**) are increasingly being adapted to permit land-based high-resolution seismic data for locating unconventional resources. This includes simultaneous source acquisition, wireless seismic technology, and an exponential growth in the number of seismic receiver channels. One major challenge to the development of high-resolution 3-D seismic is the lower data quality on land due to multiple sources of noise associated with surficial soils (Keho and Kelamis, 2012). However, land seismic surveys that use state-of-the-art acquisition and processing methods are now being used to locate tight gas trapped in complex rock systems. Such advances are leading to the first acquisition of true 3-D seismic data on land.

One unconventional production method is hydraulic fracturing (also termed 'fracking'), whereby fluids are injected into tight formations to propagate fractures that enhance the permeability of the rock, with a proppant subsequently injected to keep fractures open. The method has been extensively used in the United States for the recovery of shale oil or shale gas, although it is controversial and political due to the potential for contamination of drinking water primarily by the chemicals added to the injected fluid. Solid Earth research has relied extensively on passive monitoring of seismicity induced by plate tectonics, for example, earthquakes, and other sources of seismic energy. Using the same concepts, microseismic monitoring has now been used for over a decade for

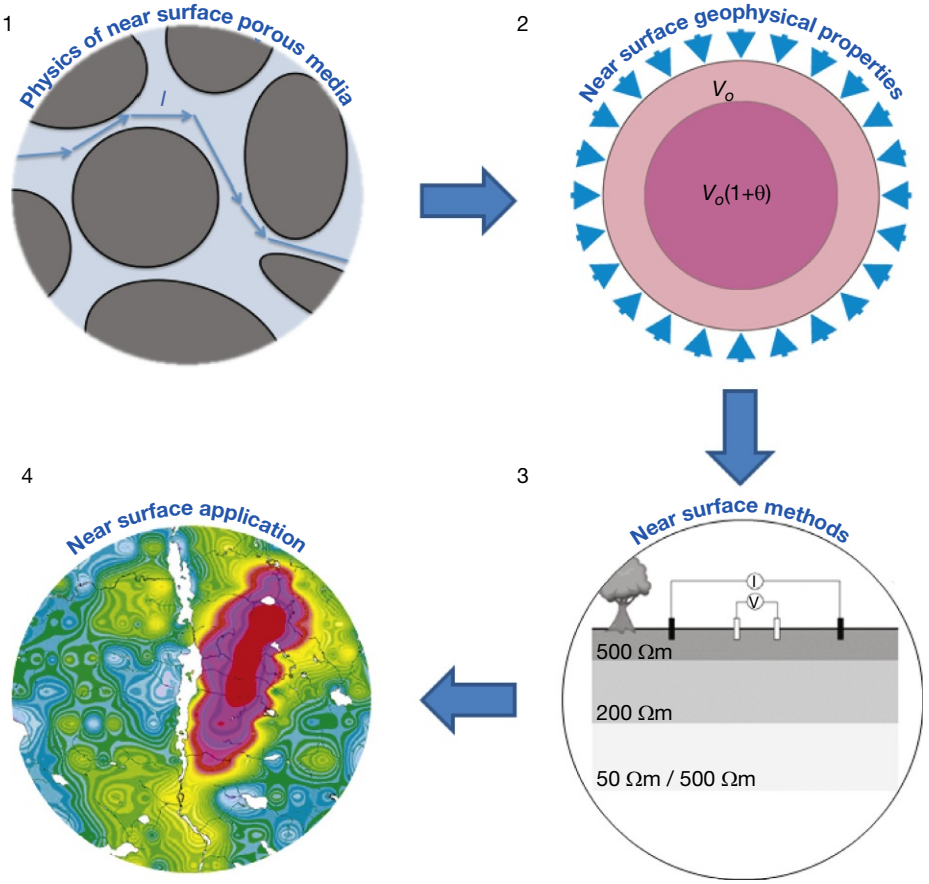


Figure 9 Overview of the organization and major sections of the volume on the near-surface Earth.

monitoring microseismicity induced by fluid injections (Maxwell and Urbancic, 2001; Maxwell et al., 2010; Sasaki, 1998). Geophones and/or other acoustic sensors are placed in wells adjacent to the injection well to record the microseismicity associated with stress changes induced by the fluid injection process. The microseismicity information is then used to map fracture growth during well stimulations.

The growing importance of unconventional fossil fuels, particularly in terms of energy dependence in the United States, is further addressed by Barth and Saiers (Chapter 11.16). Barth and Saiers also discuss the complex coupling of water resources issues to the exploitation of unconventional hydrocarbons, examining the extra water demands for extraction of unconventional, relative to conventional, hydrocarbons. They also consider the currently incomplete evidence for and against the argument that fracking can cause extensive adverse environmental impacts through the contamination of shallow potable water resources by chemicals used in the fracking process and increased dissolved methane concentrations in shallow aquifers.

Microseismicity has also been used to image deformations associated with primary production, secondary recovery, or waste injection operations associated with conventional oil (Maxwell et al., 2010). One early example of using microseismic in support of oil production is the mapping of productive reservoir fractures using downhole geophones in Clinton County, Kentucky, the United States (Rutledge et al., 1998). The overall goal of the monitoring was to delineate the reservoir fracture system. A 6-month period of monitoring was performed, during which time approximately 1800 m^3 of oil was produced from well HT1. The mechanisms were solved as composites by grouping the events from common planes. Microseismic source locations and mechanisms delineated two distinct fracture zones that formerly contained oil and were drained by the production activities and then subsequently recovered to hydrostatic pressure via brine invasion (Figure 10). This study was important as (1) it showed that low-angle fractures could be important drilling targets and (2) it provided unique information to evaluate predictions of poroelastic models regarding stress patterns around drained volumes.

11.01.10.2 Environmental Stewardship

The legacy of industrial pollution of the subsurface has highlighted the need to promote better environmental stewardship as the economies of nations continue to grow. Environmental degradation of the near-surface Earth has resulted in long-term damage to groundwater and soil resources. Subsurface pollution is a significant public health concern, particularly given the increasing dependence of nations on groundwater to meet water usage and consumption demands. Geophysical methods are increasingly being employed in attempts to characterize the extent of such degradation and monitor the effectiveness of biogeochemical processes directed at remediating contaminants. One of the most challenging

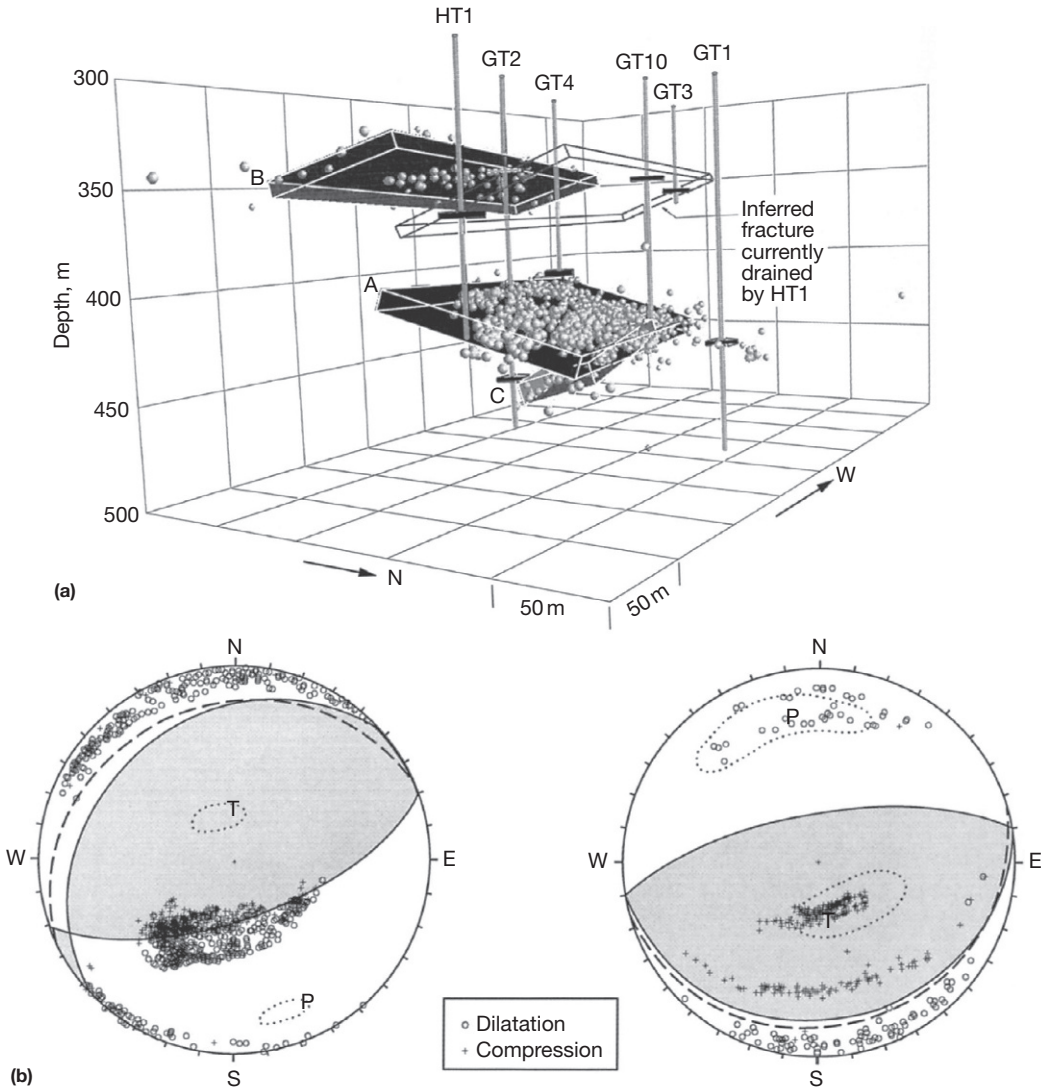


Figure 10 Results from a 6-month period of microseismic monitoring during production in Clinton County, Kentucky, the United States: (a) Perspective view of fracture planes with hypocenters within fracture volumes shown and production intervals defined by horizontal bars along wells; (b) fault-plane solutions indicating that the seismically active fractures are undergoing reverse or thrust faulting. Dashed curves on the focal hemispheres represent the orientations of the pressure (P) and tensional (T) planes determined from the respective source locations. Modified from Rutledge JT, Phillips WS, and Schuessler BK (1998) Reservoir characterization using oil-production-induced microseismicity, Clinton County, Kentucky. *Tectonophysics* 289(1): 129–152.

long-term groundwater contamination problems is a direct consequence of the legacy of nuclear weapons development and testing. In order to illustrate the contribution that geophysics is now making to such environmental legacy issues, Johnson et al. (**Chapter 11.17**) describe the extensive geophysical characterization and monitoring research performed at the US Department of Energy (DOE) Hanford Site where an unprecedented inventory of chemicals (including radionuclides) has resulted in long-term, persistent groundwater contamination that now threatens surface water bodies.

11.01.10.3 Carbon and Climate Change

NSG is now also becoming increasingly relevant to the pressing issue of carbon emissions and associated climate change. Cryosphere researchers are increasingly turning to NSG to provide information on the rates of permafrost melting and glacial ice retreat (Anandakrishnan et al., 1998; Blankenship et al., 1986; Minsley et al., 2012). Others have employed geophysical methods, including GPR and resistivity, to better understand methane production, transport, and release from peatlands (Comas et al., 2005, 2008). These important landforms contain approximately 1/3 of the terrestrial carbon sink and represent the largest natural source of methane (the second most important greenhouse gas, with a warming potential 21 times higher than CO_2) to the atmosphere. Others are turning to geophysical techniques to provide critical information on the fate of CO_2 sequestered in the subsurface in an effort to offset the ever-increasing greenhouse gas concentrations in the atmosphere.

Statoil performed the first commercial application of carbon capture and sequestration (CCS) technology at a location

in the North Sea (Sayers and Wilson, 2010). A major motivation for CCS is to stabilize CO_2 increases in the atmosphere and offset some of the pumping of CO_2 into the atmosphere from anthropogenic activities. The reliability of CCS schemes is critically dependent on technologies that can (1) provide information on the volume of CO_2 stored in the formation, (2) confirm that CO_2 is being injected at the correct locations, and (3) monitor any movement/leakage of CO_2 out of the planned storage reservoirs. Geophysical technologies have the potential to provide unique multiscale information to assess whether CCS schemes achieve adequate storage permanence, being that 99% of injected CO_2 remains securely confined after 100 years.

The majority of the efforts to monitor CCS schemes with NSG have focused on seismic monitoring due to the strong changes in acoustic velocity resulting from fluid substitution. In an early study, Harris et al. (1995) demonstrated the potential of time-lapse seismic transmission travel time tomography for monitoring a CO_2 flood pilot. Daley et al. (2007) developed a continuous active seismic cross-well monitoring system that was demonstrated to provide information on CO_2 buoyancy and associated transport during a small-scale injection. In this experiment, CO_2 injection resulted in up to 8% changes in travel time between injection and sensor wells. Passive seismic monitoring of CCS is also increasingly being explored to investigate the effect of CCS on the geomechanical properties of target reservoirs. Verdon et al. (2010) reported on 5 years of passive seismic monitoring data in response to CO_2 injection, showing evidence for a relatively low rate of seismicity associated with the CO_2 storage and minimal geomechanical deformation.

Nonseismic geophysical methods are now also increasingly being investigated for monitoring CCS schemes. Gravity and EM methods have been shown to hold potential as a lower-cost alternative to seismic for monitoring CO_2 storage as gas injection modifies both density and resistivity. Gasperikova and Hoversten (2006) performed simulations on coupled flow simulation–geophysical models representing a sandstone reservoir targeted for CO_2 storage. They showed that detectable gravity and electrical signatures could be expected based on the simulation results. The monitoring of steam injected into underground formations was an early application of the electrical resistivity tomography (ERT) technique (Ramirez et al., 1993) and the sensitivity of ERT to monitoring of CO_2 floods was addressed by Ramirez et al. (2003). In a recent study, Bergmann et al. (2012) demonstrated the use of surface–downhole resistivity tomography for imaging changes in CO_2 saturation from associated changes in electrical resistivity. In this experiment, approximately 13 500 tons of CO_2 was injected into a sandstone reservoir at an approximate depth of 650 m (**Figure 11**). Technical challenges of acquiring reliable ERT data for this application, including a sufficiently large power source and signal-to-noise ratio at the borehole sensors (electrodes), were addressed in this study. Resistivity images of the subsurface revealed preferential migration of CO_2 away from the injection location due to the transport characteristics of the aquifer. Estimates of the CO_2 saturation were generated from the resistivity index, that is, the ratio of resistivity at time t after injection to the resistivity before injection. Resulting images were interpreted in terms of buoyancy-driven displacement of brines within the formation.

11.01.11 Future Needs in NSG

We finish by considering the most pressing research needs to further our geophysical understanding of the near-surface Earth. Firstly, we note that the present understanding of the link between geophysical properties and the physicochemical properties of the Earth (commonly known as petrophysics due to the origin of this field in petroleum exploration) is limited in that these relations are generally considered static, that is, they consider physicochemical properties of rocks that are primarily determined by geologic timescales. Yet, as we have discussed, the near surface is dynamic in that biogeochemical transformations modify both the physicochemical and geophysical properties of Earth materials over very short timescales (days to thousands of years) relative to geologic time. One example is the biomineralization process driven by microbial-mediated sulfate reduction in the presence of dissolved iron shown in **Figure 8**. Existing petrophysical relations cannot model such processes. Consequently, we foresee a need for a new era of research on time-dependent petrophysics that could be used to better understand the geophysical signatures resulting from natural biogeochemical processes and accelerated modifications due to human alteration of the subsurface.

Coupled to this basic research, we foresee a new era of long-term geophysical observations of the near-surface Earth directed at environmental stewardship. Here, an analogy to the field of medicine is justified, where geophysical sensing has long been used to monitor the health of the human body from observations made at its surface. Geophysical sensing of the near surface is increasingly needed to monitor the health of the Earth. Examples covered in this chapter include monitoring of (1) accelerated permafrost thawing and postulated increasing methane emissions due to anthropogenic-induced climate warming, (2) contaminant transport and attenuation/biodegradation (engineered/accelerated or long-term natural), and (3) hydrocarbon and groundwater reservoir depletion due to continued exploitation of resources. Examples of improved understanding of subsurface processes resulting from geophysical monitoring are now widespread and include (1) the effectiveness of enhanced oil recovery processes (Greaves and Fulp, 1987), (2) hydrogeologic controls on focused groundwater discharge into rivers (Johnson et al., 2012), (3) coupled microbial and geochemical processes associated with enhanced attenuation of contaminants (Williams et al., 2009), and (4) vegetation-regulated soil moisture dynamics (Robinson et al., 2012). Improved monitoring capabilities require the development of new geophysical technologies capable of long-term, automated data acquisition, often in remote environments. The benefits of long-term geophysical monitoring are demonstrated by Hilbich et al. (2008) in a 7-year study where resistivity monitoring was used to monitor freezing and thawing cycles associated with permafrost evolution. Geophysical instrumentation is required that overcomes the current inflexibility of instruments designed primarily for traditional exploration geophysics when implemented for monitoring. Autonomous geophysical monitoring platforms should ultimately be developed to provide noninvasive proxy measurements of chemical, physical, and biological processes operating over long timescales.

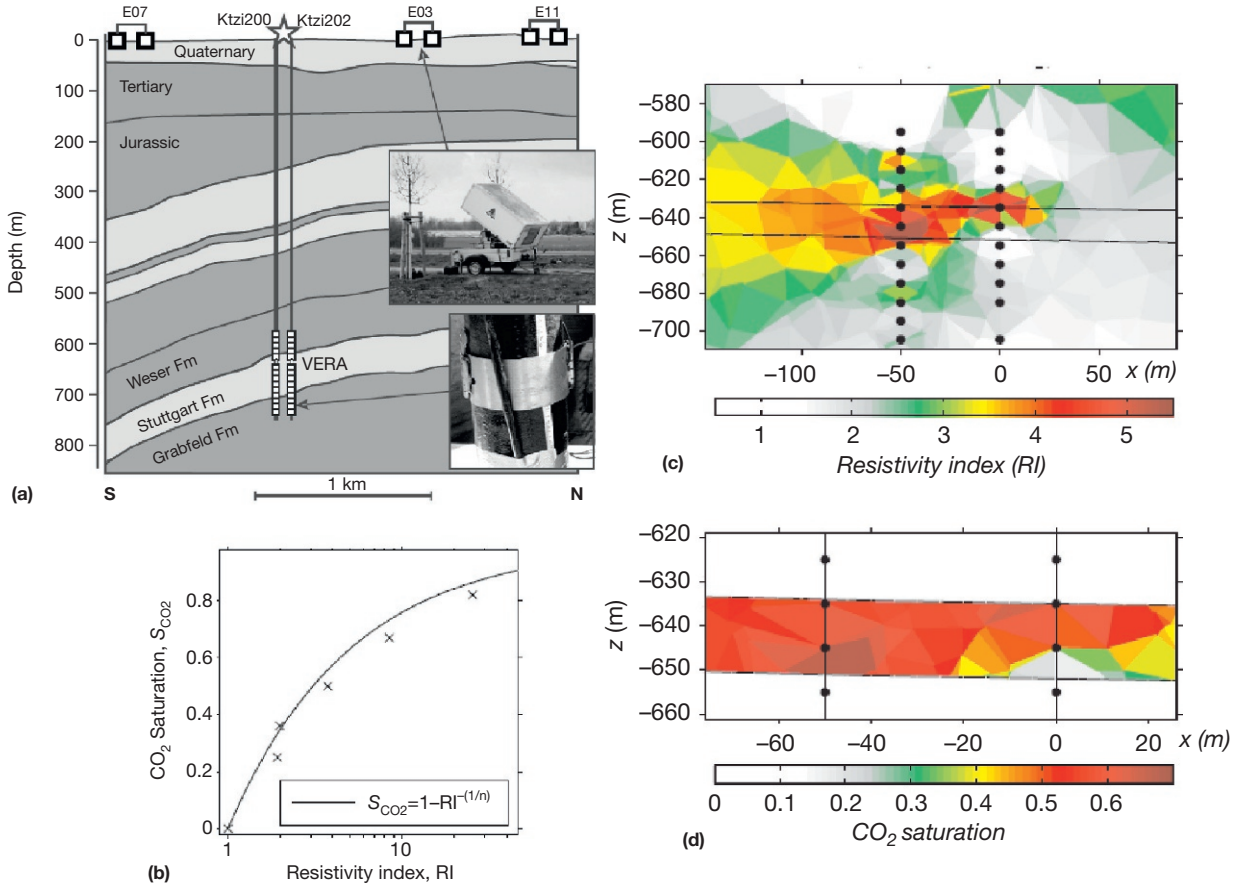


Figure 11 Surface–borehole resistivity monitoring of CO_2 injection into the Stuttgart Formation at the pilot storage in Ketzin, Germany (Bergmann et al., 2012): (a) 2-D section of the experimental setup showing power source (top inset) supplying surface current electrodes and borehole electrodes used for recording voltages in the Stuttgart Formation (bottom inset); (b) petrophysical relationship for Stuttgart Formation used to convert resistivity index (RI) to CO_2 saturation (S_{CO_2}). Reproduced from Kummerow J and Spangenberg E (2011) Experimental evaluation of the impact of the interactions of CO_2–SO_2, brine, and reservoir rock on petrophysical properties: A case study from the Ketzin test site, Germany. *Geochemistry, Geophysics, Geosystems* 12(5). (c) Resistivity index (RI) showing increases in resistivity after 13.5 kT CO_2 injected; (d) resulting distribution of CO_2 saturation calculated from RI image shown in (c) using petrophysical relationship shown in (b). Figures (a), (c), and (d) modified from Bergmann P, Schmidt-Hattenberger C, Kiessling D, et al. (2012) Surface-downhole electrical resistivity tomography applied to monitoring of CO_2 storage at Ketzin, Germany. *Geophysics* 77(6): B253–B267.

One example is the development of electrical geophysical monitoring systems for long-term autonomous monitoring of vulnerability of coastal freshwater aquifers to saline intrusion. The British Geological Survey recently developed the automated time-lapse electrical resistivity (ALERT) system (Ogilvy et al., 2009) to perform such monitoring of coastal dynamics using electrical resistivity methods as reviewed by Binley (**Chapter 11.08**). ALERT has been employed in Europe at critical coastal locations suffering from saltwater intrusion into aquifers. One example is the study of the River Andarax, Almería, Spain, to monitor the seawater–freshwater interface in the underlying Quaternary aquifer using an array of electrodes extending 1.6 km upstream from the shoreline (**Figure 12**). The unmanned ALERT system generates a daily image of the changes in resistivity due to movement of the freshwater–saltwater interface driven by pumping and irrigation schemes impacting the aquifer. This monitoring system provides unprecedented information on saline intrusion in this region over spatial scales that could never be practically investigated using boreholes. The ALERT system has also been deployed to monitor slow mass-movement processes on hillslopes (Chambers et al., 2009).

In a logical extension of traditional airborne geophysics, we anticipate future deployments of geophysical sensors from unmanned aircraft systems (UASs) or drones. Such airborne platforms, primarily developed for surveillance, could be used to conduct geophysical surveys over remote areas at unprecedented spatial scales. Very recent studies demonstrate this potential by using drones equipped with thermal cameras and gas sensors to survey active volcanoes (Williams, 2013) where surveying such regions in manned aircraft is high risk. Currently, very few geophysical sensors can be miniaturized to sizes and payload limitations of mini UAS. The magnetometer and gamma-ray spectrometer are two exceptions that are currently deployable from UAS. Stoll and Moritz (2013) reported on the deployment of a fluxgate

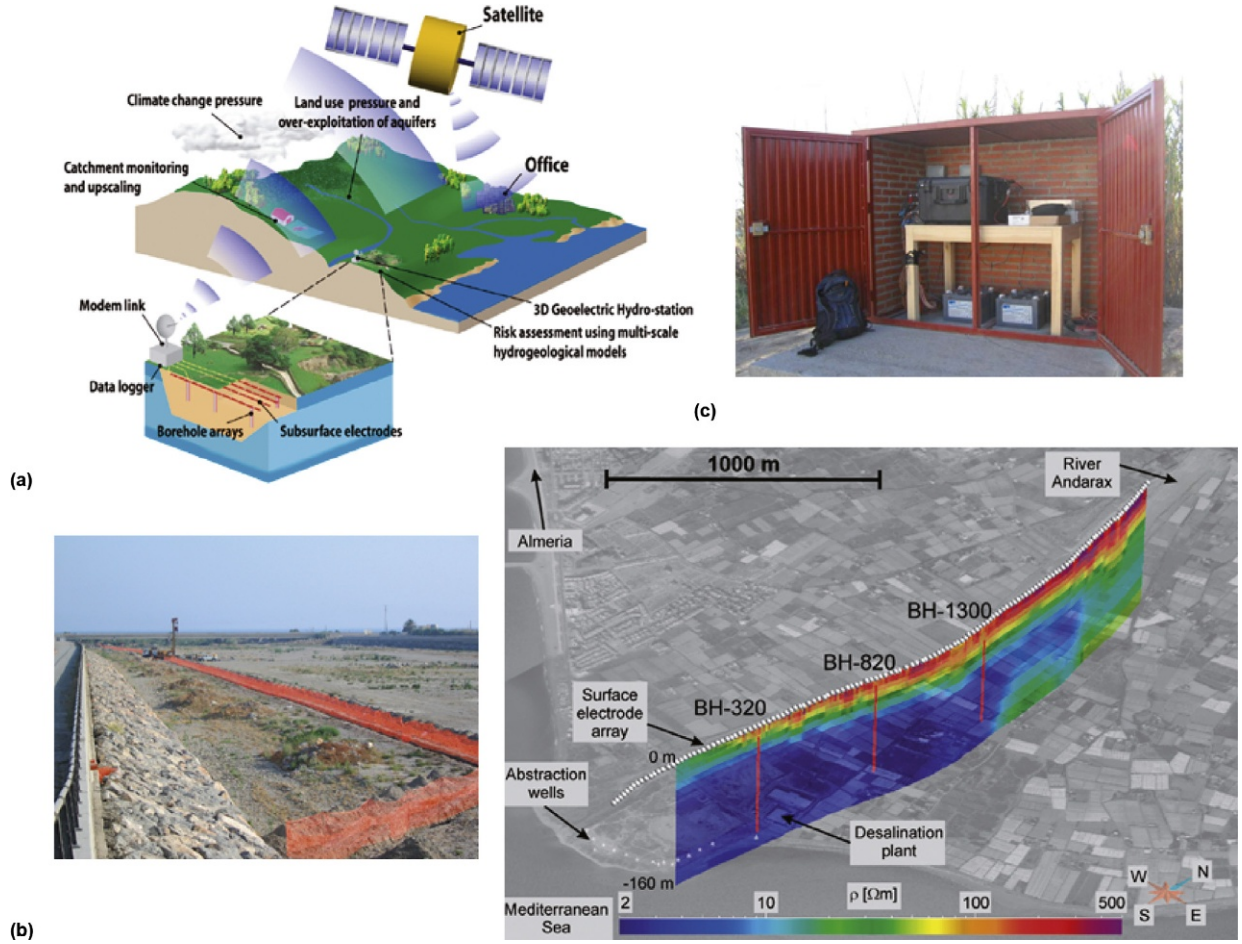


Figure 12 Concept of electrical resistivity monitoring of saline intrusion as demonstrated by the British Geological Survey ALERT system: (a) Schematic concept of autonomous resistivity monitoring setup; (b) installation of sensors (electrodes) at a site impacted by saline intrusion in Almería, Spain; (c) typical autonomous monitoring instrumentation; (d) aerial photograph with 2-D ALERT image overlay. Visualization of the subsurface aquifer showing conductive saline water (blue), resistive freshwater (red), and a mixing zone (green). Reproduced from Ogilvy RD, Meldrum PI, Kuras O, et al. (2009) Automated monitoring of coastal aquifers with electrical resistivity tomography. *Near Surface Geophysics* 7(5–6): 367–375.

magnetometer mounted on a UAS and demonstrated its capability for locating ferrous objects. Drones might be used to map the distribution of oil contamination in remote coastal marshes and wetlands following disasters such as the BP Deepwater Horizon oil spill. Mapping of UXO in war zones is an obvious important application of geophysical sensors from drones.

It is increasingly recognized that the development of joint (and coupled) inversion frameworks has great potential for improving the information content on the physicochemical properties of the near-surface Earth that is extractable from geophysical datasets. For example, the monitoring of CCS is likely to advance when seismic data are combined with EM and gravity datasets and the synergies of jointly interpreting the data are realized. Seismic reflection data can provide high-resolution qualitative information on geologic structures; in contrast, EM and gravity are low-resolution geophysical methods but potentially contain quantitative information on volume and mass percent of CO_2 in the subsurface (Lumley, 2010). Similar synergistic benefits exist for joint inversion of geophysical datasets for monitoring reservoir production activities and hydrogeologic processes. In the latter case, geophysical imaging data that provide spatially extensive and continuous information of proxies of flow and transport parameters are integrated with sensors that provide direct measurements of flow and transport parameters at a limited number of specific locations.

In summary, geophysics is making an increasingly important contribution to our understanding of the extractable resources of the Earth, the impacts of human degradation of the subsurface, and the critical hydrogeologic and biogeochemical processes within the shell of the Earth that is subject to human interaction. This outer shell is also characterized by the presence of fluids in open, connected pores that exert a profound effect on geophysical properties. Advances in instrumentation and data acquisition platforms are providing ever-increasing amounts of data from sensors distributed over multiple spatial scales. The continued development of geophysical instrumentation for autonomous time-lapse monitoring is providing opportunities to observe dynamic near-surface processes with unique spatiotemporal resolution. However, perhaps the greatest opportunities for improving the understanding of near-surface processes will come

from interdisciplinary studies where geophysics is performed in conjunction with geochemical and biogeochemical studies and the resulting data are interpreted in unison. Adopting such an approach, geophysics can be fully exploited for the unique non-invasive information it provides at large spatial scales.

References

Allred B, Daniels JJ, and Ehsani MR (eds.) (2008) *Handbook of Agricultural Geophysics*. Boca Raton, FL: CRC Press.

Anandakrishnan S, Blankenship DD, Alley RB, and Stoffa PL (1998) Influence of subglacial geology on the position of a West Antarctic ice stream from seismic observations. *Nature* 394(6688): 62–65.

Atekwana EA and Slater LD (2009) Biogeophysics: A new frontier in Earth science research. *Reviews of Geophysics* 47(4), RG4004.

Bergmann P, Schmidt-Hattenberger C, Kiessling D, et al. (2012) Surface-downhole electrical resistivity tomography applied to monitoring of CO_2 storage at Ketzin, Germany. *Geophysics* 77(6): B253–B267.

Berthling I, Etzelmüller B, Isaksen K, and Sollid JL (2000) Rock glaciers on Prins Karls Forland. II: GPR soundings and the development of internal structures. *Permafrost and Periglacial Processes* 11(4): 357–369.

Blankenship DD, Bentley CR, Rooney ST, and Alley RB (1986) Seismic measurements reveal a saturated porous layer beneath an active Antarctic ice stream. *Nature* 322: 54–57.

Brantley S, Goldhaber MB, and Ragnarsdottir KV (2007) Crossing disciplines and scales to understand the. *Critical Zone: Elements* 3(5): 307–314.

Brantley S, White T, White A, et al. (2006) Frontiers in exploration of the critical zone. In: *Report of a Workshop Sponsored by the National Science Foundation (NSF), October 30, pp. 24–26, 2005, Newark, DE*.

Chambers JE, Meldrum PI, Gunn DA, et al. (2009) Hydrogeophysical monitoring of landslide processes using automated time-lapse electrical resistivity tomography (ALERT). In: *15th European Meeting of Environmental and Engineering Geophysics, September, 2009*.

Chew KJ (2014) The future of oil: Unconventional fossil fuels. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 372(2006): 20120324.

Chorover J, Kretzschmar R, Garcia-Pichel F, and Sparks DL (2007) Soil biogeochemical processes within the critical zone. *Elements* 3(5): 321–326.

Chorover J, Troch PA, Rasmussen C, et al. (2011) How water, carbon, and energy drive landscape evolution and surface water dynamics: The Jemez River Basin – Santa Catalina Mountains Critical Zone Observatory. *Vadose Zone Journal* 10: 884–889 (special issue on the Critical Zone).

Christensen TR, Johansson T, Åkerman HJ, et al. (2004) Thawing sub-arctic permafrost: Effects on vegetation and methane emissions. *Geophysical Research Letters* 31(4).

Clay J (2004) *World Agriculture and the Environment: A Commodity-by-Commodity Guide to Impacts and Practices*. Washington, DC: Island Press.

Colombo D, Cogan M, Hallinan S, Mantovani M, Virgilio M, and Soyer W (2008) Near-surface P-velocity modelling by integrated seismic, EM, and gravity data: Examples from the Middle East. *First Break* 26(10).

Comas X, Slater L, and Reeve A (2005) Spatial variability in biogenic gas accumulations in peat soils is revealed by ground penetrating radar (GPR). *Geophysical Research Letters* 32(8).

Comas X, Slater L, and Reeve A (2008) Seasonal geophysical monitoring of biogenic gases in a northern peatland: Implications for temporal and spatial variability in free phase gas production rates. *Journal of Geophysical Research: Biogeosciences (2005–2012)* 113(G1).

Daley TM, Solbau RD, Ajo-Franklin JB, and Benson SM (2007) Continuous active-source seismic monitoring of CO_2 injection in a brine aquifer. *Geophysics* 72(5): A57–A61.

Davis K, Li Y, and Batzle M (2008) Time-lapse gravity monitoring: A systematic 4D approach with application to aquifer storage and recovery. *Geophysics* 73(6): WA61–WA69.

Detheir D and Lazarus E (2006) Geomorphic inferences from regolith thickness, chemical denudation and CRN erosion rates near the glacial limit, Boulder Creek catchment and vicinity, Colorado. *Geomorphology* 75(3–4): 384–399. http://dx.doi.org/10.1016/j.geomorph.2005.07.029.

Dietrich W (2010) Bedrock matters. *Geochimica et Cosmochimica Acta* 74(11 Suppl. 1): A232.

d'Ozouville N, Auken E, Sorensen K, et al. (2008) Extensive perched aquifer and structural implications revealed by 3D resistivity mapping in a Galapagos volcano. *Earth and Planetary Science Letters* 269(3): 518–522.

Flores Orozco A, Williams KH, Long PE, Hubbard SS, and Kemna A (2011) Using complex resistivity imaging to infer biogeochemical processes associated with bioremediation of an uranium-contaminated aquifer. *Journal of Geophysical Research: Biogeosciences (2005–2012)* 116(G3).

Gasperikova E and Hoversten GM (2006) A feasibility study of nonseismic geophysical methods for monitoring geologic CO_2 sequestration. *The Leading Edge* 25(10): 1282–1288.

Gleick PH and Palaniappan M (2010) Peak water limits to freshwater withdrawal and use. *Proceedings of the National Academy of Sciences* 107(25): 11155–11162.

Graham R, Rossi A, and Hubbert R (2010) Rock to regolith conversion: Producing hospitable substrates for terrestrial ecosystems. *GSA Today* 4–9. http://dx.doi.org/10.1130/GSAT57A.

Greaves RJ and Fulp TJ (1987) Three-dimensional seismic monitoring of an enhanced oil recovery process. *Geophysics* 52(9): 1175–1187.

Harris JM, Nolen-Hoeksema RC, Langan RT, Van Schaack M, Lazaratos SK, and Rector JW III (1995) High-resolution crosswell imaging of a west Texas carbonate reservoir: Part 1-Project summary and interpretation. *Geophysics* 60(3): 667–681.

Hauck C and MühII DV (2003) Inversion and interpretation of two-dimensional geoelectrical measurements for detecting permafrost in mountainous regions. *Permafrost and Periglacial Processes* 14: 305–318. http://dx.doi.org/10.1002/ppp.462.

Hilbich C, Hauck C, Hoelzle M, et al. (2008) Monitoring mountain permafrost evolution using electrical resistivity tomography: A 7-year study of seasonal, annual, and long-term variations at Schilthorn, Swiss Alps. *Journal of Geophysical Research* 113: F01S90.

Hinkel KM, Doolittle JA, Bockheim JG, et al. (2001) Detection of subsurface permafrost features with ground-penetrating radar, Barrow, Alaska. *Permafrost and Periglacial Processes* 12: 179–190. http://dx.doi.org/10.1002/ppp.369.

Huang H and Won IJ (2003) Characterization of UXO-like targets using broadband electromagnetic induction sensors. *IEEE Transactions on Geoscience and Remote Sensing* 41(3): 652–663.

Johnson TC, Slater LD, Ntarlagiannis D, Day-Lewis FD, and Elwaseif M (2012) Monitoring groundwater-surface water interaction using time-series and time-frequency analysis of transient three-dimensional electrical resistivity changes. *Water Resources Research* 48(7).

Keho TH and Kelamis PG (2012) Focus on land seismic technology: The near-surface challenge. *The Leading Edge* 31(1): 62–68.

Kulessa B, Hubbard B, and Brown GH (2003) Cross-coupled flow modeling of coincident streaming and electrochemical potentials and application to subglacial self-potential data. *Journal of Geophysical Research* 108(B8): 2381.

Kummerow J and Spangenberg E (2011) Experimental evaluation of the impact of the interactions of CO_2–SO_2, brine, and reservoir rock on petrophysical properties: A case study from the Ketzin test site, Germany. *Geochemistry, Geophysics, Geosystems* 12(5).

Lumley D (2010) 4D seismic monitoring of CO_2 sequestration. *The Leading Edge* 29(2): 150–155.

Lunt IA, Hubbard SS, and Rubin Y (2005) Soil moisture content estimation using ground-penetrating radar reflection data. *Journal of Hydrology* 307(1): 254–269.

Machguth H, Eisen O, Paul F, and Hoelzle M (2006) Strong spatial variability of snow accumulation observed with helicopter-borne GPR on two adjacent Alpine glaciers. *Geophysical Research Letters* 33: L13503.

Maxwell SC, Rutledge J, Jones R, and Fehler M (2010) Petroleum reservoir characterization using downhole microseismic monitoring. *Geophysics* 75(5): 75A129–75A137.

Maxwell SC and Urbancic TI (2001) The role of passive microseismic monitoring in the instrumented oil field. *The Leading Edge* 20(6): 636–639.

Mekonnen MM and Hoekstra AY (2011) The green, blue and grey water footprint of crops and derived crop products. *Hydrology and Earth System Sciences* 15(5): 1577–1600.

Minsley BJ, Abraham JD, Smith BD, et al. (2012) Airborne electromagnetic imaging of discontinuous permafrost. *Geophysical Research Letters* 39: L02503.

Murray T, Stuart GW, Fry M, Gamble NH, and Crabtree MD (2000) Englacial water distribution in a temperate glacier from surface and borehole radar velocity analysis. *Journal of Glaciology* 46(154): 389–398.

Musil M, Maurer H, Green A, et al. (2002) Shallow seismic surveying of an Alpine rock glacier. *Geophysics* 67(6): 1701–1710.

National Research Council (2001) *Basic Research Opportunities in Earth Science*. Washington, DC: The National Academies Press.

Ogilvy RD, Meldrum PI, Kuras O, et al. (2009) Automated monitoring of coastal aquifers with electrical resistivity tomography. *Near Surface Geophysics* 7(5–6): 367–375.

O'Neel S, Marshall HP, McNamara DE, and Pfeffer WT (2007) Seismic detection and analysis of icequakes at Columbia Glacier, Alaska. *Journal of Geophysical Research: Earth Surface (2003–2012)* 112(F3).

Pan L, Adamchuk VI, Martin DL, Schroeder MA, and Ferguson RB (2013) Analysis of soil water availability by integrating spatial and temporal sensor-based data. *Precision Agriculture* 14(4): 1–20.

Papadopoulos NG, Tsourlos P, Tsokas GN, and Sarris A (2006) Two-dimensional and three-dimensional resistivity imaging in archaeological site investigation. *Archaeological Prospection* 13(3): 163–181.

Parsekian AD, Grosse G, Walbrecker JO, et al. (2013) Detecting unfrozen sediments below thermokarst lakes with surface nuclear magnetic resonance. *Geophysical Research Letters* 40: 535–540. http://dx.doi.org/10.1002/grl.50137.

Pasion LR and Oldenburg DW (2001) A discrimination algorithm for UXO using time domain electromagnetics. *Journal of Environmental & Engineering Geophysics* 6(2): 91–102.

Postel SL, Daily GC, and Ehrlich PR (1996) Human appropriation of renewable fresh water. *Science* 271(5250): 785–787.

Qamar A (1988) Calving icebergs: A source of low-frequency seismic signals from Columbia Glacier, Alaska. *Journal of Geophysical Research* 93(B6): 6615–6623.

Ramirez A, Daily W, LaBrecque D, Owen E, and Chesnut D (1993) Monitoring an underground steam injection process using electrical resistance tomography. *Water Resources Research* 29(1): 73–87.

Ramirez E, Francou B, Ribstein P, et al. (2001) Small glaciers disappearing in the tropical Andes: A case-study in Bolivia: Glaciar Chacaltaya (16 S). *Journal of Glaciology* 47(157): 187–194.

Ramirez AL, Newmark RL, and Daily WD (2003) Monitoring carbon dioxide floods using electrical resistance tomography (ERT): Sensitivity studies. *Journal of Environmental & Engineering Geophysics* 8(3): 187–208.

Robinson JL, Slater LD, and Schäfer KV (2012) Evidence for spatial variability in hydraulic redistribution within an oak–pine forest from resistivity imaging. *Journal of Hydrology* 430: 69–79.

Rubin Y and Hubbard SS (eds.) (2005) In: *Hydrogeophysics*, vol. 50. Dordrecht, The Netherlands: Springer.

Rutledge JT, Phillips WS, and Schuessler BK (1998) Reservoir characterization using oil-production-induced microseismicity, Clinton County, Kentucky. *Tectonophysics* 289(1): 129–152.

Sasaki S (1998) Characteristics of microseismic events induced during hydraulic fracturing experiments at the Hijiori hot dry rock geothermal energy site, Yamagata, Japan. *Tectonophysics* 289(1): 171–188.

Sayers C and Wilson T (2010) An introduction to this special section: CO_2 sequestration. *The Leading Edge* 29(2): 148–149.

Schwamborn G, Andreev A, Rachold V, et al. (2002a) Evolution of Lake Nikolay, Arga Island, Western Lena River delta, during Late Pleistocene and Holocene time. *Polarforschung* 70: 69–82.

Schwamborn GJ, Dix JK, Bull JM, and Rachold V (2002b) High-resolution seismic and ground penetrating radar – Geophysical profiling of a thermokarst lake in the western Lena Delta, Northern Siberia. *Permafrost and Periglacial Processes* 13: 259–269.

Scollar I, Tabbagh A, Hesse A, and Herzog I (1990) *Archaeological Prospecting and Remote Sensing*. Cambridge: Cambridge University Press.

Slater L and Atekwana E (2013) Geophysical signatures of subsurface microbial processes. *Eos, Transactions American Geophysical Union* 94(8): 77–78.

Slater L, Knight R, Singha K, Binley A, and Atekwana E (2006) Near-surface geophysics: A new focus group. *Eos, Transactions American Geophysical Union* 87(25): 249.

Slater L, Ntarlagiannis D, Personna YR, and Hubbard S (2007) Pore-scale spectral induced polarization signatures associated with FeS biomineral transformations. *Geophysical Research Letters* 34(21): L21404.

Stoll J and Moritz D (2013) Unmanned aircraft systems for rapid near surface geophysical measurements. In *75th EAGE Conference & Exhibition-Workshops*.

Tarantola A (1984) Inversion of seismic reflection data in the acoustic approximation. *Geophysics* 49(8): 1259–1266. http://dx.doi.org/10.1190/1.1441754.

Verdon JP, Kendall JM, White DJ, Angus DA, Fisher QJ, and Urbancic T (2010) Passive seismic monitoring of carbon dioxide storage at Weyburn. *The Leading Edge* 29(2): 200–206.

Walter KM, Zimov SA, Chanton JP, Verbyla D, and Chapin FS (2006) Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming. *Nature* 443(7107): 71–75.

West JJ and Plug LJ (2008) Time-dependent morphology of thaw lakes and taliks in deep and shallow ground ice. *Journal of Geophysical Research* 113: F01009.

Weymouth JW (1986) Archaeological site surveying program at the University of Nebraska. *Geophysics* 51(3): 538–552.

Williams SC (2013) Studying volcanic eruptions with aerial drones. *Proceedings of the National Academy of Sciences* 110(27): 10881.

Williams KH, Kemna A, Wilkins MJ, et al. (2009) Geophysical monitoring of coupled microbial and geochemical processes during stimulated subsurface bioremediation. *Environmental Science & Technology* 43(17): 6717–6723.

Winship P, Binley A, and Gomez D (2006) Flow and transport in the unsaturated Sherwood sandstone: Characterization using cross-borehole geophysical methods. *Geological Society, London, Special Publications* 263(1): 219–231.

11.02 Physics of Porous Media: Fluid Flow Through Porous Media

Y Bernabé, Massachusetts Institute of Technology, Cambridge, MA, USA
A Maineult, Institut de Physique du Globe de Paris, Paris, France

© 2015 Elsevier B.V. All rights reserved.

11.02.1	**Introduction**	19
11.02.2	**Materials**	20
11.02.3	**Microstructure**	21
11.02.3.1	Geometry of the Pore Space	21
11.02.3.2	Topology	22
11.02.3.3	Summary	23
11.02.4	**Single-Phase Flow**	24
11.02.4.1	Darcy's Law, Permeability	24
11.02.4.2	Measurement of Permeability	25
11.02.4.3	Permeability Models	26
11.02.4.4	Summary	28
11.02.5	**Solute Transport**	28
11.02.5.1	General Considerations	28
11.02.5.2	Taylor Dispersion	29
11.02.6	**Multiphase Flow**	30
11.02.6.1	Wettability, Capillary Pressure, Drainage, and Imbibition	30
11.02.6.2	Two-Phase Flow, Relative Permeabilities	31
11.02.7	**Processes Affecting Permeability**	32
11.02.7.1	Examples of Process-Specific PPRs	33
11.02.7.1.1	Elastic compression	33
11.02.7.1.2	Nonelastic compression	34
11.02.7.1.3	Dissolution–precipitation, alteration, and chemical reactions	35
11.02.7.2	Natural Processes, Diagenesis	36
11.02.7.3	Summary	37
11.02.8	**Concluding Statement**	37
Acknowledgments		37
References		37

11.02.1 Introduction

The scientific and economical importance of porous media in our society is hard to overstate. Porous media are ubiquitous around us. Construction materials such as wood, concrete, or bricks; engineering materials such as ceramics; and even the very paper on which this book is printed are porous (for more examples of porous media, see Section 1.3 of Dullien, 1992). And, of course, porosity controls many useful properties of these materials, allowing them to be used as thermal insulators, filters, fluid absorbents, and so forth. Geologic materials near the surface of the Earth are also porous, and many natural processes, from the absorption of nutrients in topsoil by plants to the accumulation of hydrocarbons in geologic reservoirs, depend on the presence of connected pore space and fluids permeating it. Modeling the physical properties of porous geologic materials is therefore an important subfield of geophysics and will be the subject of this chapter (see, e.g., Berryman, 1995; Guéguen and Palciauskas, 1994; Mavko et al., 2003, or Zinszner and Pellerin, 2007, for previously published monographs on this theme). In order to limit the size of this presentation and to avoid excessive overlap with the other chapters, we will concentrate on permeability and various aspects of fluid flow through porous rocks. The basic concepts involved (e.g., microstructure, heterogeneity, connectivity, and upscaling) are generally relevant for modeling other rock properties.

The ultimate goal of rock physics is to construct an accurate description of the phenomena of interest at macroscopic scales. There are many ways to accomplish this, including establishing the macroscopic equations directly based on heuristic arguments (e.g., Biot, 1962). However, the most rigorous approach is to start with a pore-scale description and apply one of many upscaling methods to the microscopic equations (e.g., volume-averaging theorem (Burridge and Keller, 1981; Mei and Auriault, 1989; Pride et al., 1992; Revil and Leroy, 2004; Revil and Linde, 2006; Slattery, 1967; Whitaker, 1999), homogenization (Auriault, 2002; Bourgeat, 1997; Noetinger et al., 2001; Panfilov, 2000), effective medium theory (Choy, 1999), percolation theory (Hunt and Ewing, 2009; Kirkpatrick, 1973), and renormalization group (Hristopulos, 2003; King, 1989); see also reviews by Berryman, 2005; Dagan, 1989; Farmer, 2002; Gelhar, 1993; Renard and de Marsily, 1997; Vereecken et al., 2007)). The form of the macroscopic equations may or may not be identical to the microscopic ones. Sometimes, entirely new

phenomena arise at macroscopic scales (e.g., porous rocks are often observed to be nonlinear elastic, while their mineral constituents are all linear elastic; this behavior could be caused, among other things, by the gradual closure of crack-like pores with different aspect ratios).

Although the chemical composition and physical properties of the solid constituents of rocks can play a significant role in natural processes such as the compaction of sediments with depth or in physical phenomena like the transmission of elastic waves through porous media, the most important factor affecting the physical properties of porous rocks, in general, and the transport properties, in particular, is the geometry of the pore space (e.g., Dullien, 1992). Early attempts to characterize the three-dimensional (3D) pore geometry of rocks used serial sectioning, a sample destructing technique with a very low resolution in the normal direction (e.g., DeHoff et al., 1972; Dullien, 1992; MacDonald et al., 1986; Scott et al., 1988; Vogel, 1997). Today, researchers are making immense progress on a range of nondestructive, high-resolution 3D imaging techniques (x-ray microtomography (Al-Raoush and Willson, 2005; Auzerais et al., 1996; Gaillard et al., 2007; Lindquist et al., 2000; Spanne et al., 1994; Wildenschild et al., 2002), laser scanning confocal microscopy (Fredrich, 1999; Fredrich et al., 1995; Menéndez et al., 2001; Petford et al., 1999, 2001), and nuclear magnetic resonance microscopy (Baldwin et al., 1996; Doughty and Tomutsa, 1997)). These new technological advances open the possibility of fast, high-resolution determination of the 3D pore geometry in many types of porous media. This development, in conjunction with steady improvement of 3D numerical methods, has the potential to revolutionize rock physics because it is now becoming possible to simulate complex pore-scale processes numerically (finite differences, finite elements, and lattice-gas and lattice Boltzmann methods, e.g., Adler et al., 1990; Aharonov and Rothman, 1993; Arns et al., 2001, 2002, 2004a; Bentz and Martys, 1994; Bosl et al., 1998; Ferreol and Rothman, 1995; Gunstensen and Rothman, 1993; Kameda et al., 2006; Keehm et al., 2001; Manz et al., 1999; Øren and Bakke, 2002; Papafotiou et al., 2008; Poutet et al., 1996; Rothman, 1990; see also the review by Sahimi, 1995). Mention should also be made of methods in which the 3D pore structure is idealized into a network of channels, allowing a much simpler estimation of the single- or multiphase flow properties of porous media (e.g., Arns et al., 2004b; Blunt et al., 2002; Sok et al., 2002; Vogel and Roth, 2001). Numerical simulations such as those mentioned earlier, if run over volumes containing a sufficiently large number of pores, may allow verifying presently accepted (or devising new) macroscopic theories of the phenomena under consideration.

This chapter is organized as follows: in **Section 11.02.2**, we will give a brief description of the geomaterials most commonly found near the surface of the Earth; **Section 11.02.3** will be devoted to the geometric and topological characterization of the pore space; **Section 11.02.4** will review the basic aspects of the physics of flow through porous media; solute transport will be succinctly treated in **Section 11.02.5** and multiphase flow in **Section 11.02.6**; and we will finally consider the evolution of permeability in geomaterials subjected to various mechanical or chemical processes in **Section 11.02.7**.

11.02.2 Materials

Many different types of geomaterials can be found at or near the surface of the Earth. The classification of these materials is difficult, owing to the vast differences in their mineralogy, structure, or genetic processes. As far as environmental problems are concerned, the most important class of materials is the sedimentary group, encompassing all the variations from soils and unconsolidated sediments to indurated sedimentary rocks (e.g., Folk, 1959; Mitchell and Soga, 2005; Morse, 2003; Pettijohn, 1975; Pettijohn et al., 1987; Wilson, 1999).

Sediments are granular materials produced in various ways (e.g., weathering of preexisting rocks, and volcanic eruptions), transported by wind or water, and deposited in low topographic regions. They are usually denominated according to their typical particulate size. From coarse to fine, geologists distinguish gravels (2 to 256 mm), sands (1/16 to 2 mm), silts (1/256 to 1/16 mm), and clays (less than 1/256 mm). Of course, the materials in each category may contain a minor portion of greatly oversized or undersized particles. Because of their large size, gravels primarily consist of polymineralic (possibly porous) rock fragments. Sands and silts generally contain a majority of monomineralic grains. Their mineralogical composition can vary widely, depending on many factors such as the source rocks or the remobilization and mixing of different sediments. Silicate, carbonate, and volcanic sands are commonly observed all over the globe. Clays correspond to a specific mineralogical group, namely, hydrous aluminosilicates. They have a characteristic sheet structure of negatively charged layers bound together by hydrated cations, which allows them to adsorb water between the layers (leading, in the case of smectite, to the phenomenon known as clay swelling, e.g., Bleam, 1993).

In loose sediments, the grains are in tangential contact, forming a (somewhat unstable) force-transmitting structure and leaving an open network of pores between them. The stress field in individual grains is extremely nonuniform (high near the points of contact and small elsewhere) and is controlled by the shape of the grains and grain contacts. The distribution of contact forces in the granular framework is also highly nonuniform and depends on the packing (itself a function of, e.g., grain size and shape and sorting). The pore space is complementary to the grain volume and its characteristics (i.e., pore size, pore shape, and topology) must therefore depend on the same set of factors as the solid framework.

During burial, sediments become lithified by diagenesis, that is, a set of mechanical and chemical processes occurring at various depths. In broad terms, diagenesis tends to reduce porosity, although some porosity-enhancing processes such as dissolution of labile minerals can be activated in some situations (e.g., Bjørlykke et al., 1989; Choquette and Pray, 1970; Dunoyer de Segonzac, 1970; Morse, 2003; Pettijohn et al., 1987; Pittman, 1979). In sandstones, porosity is mostly removed by cementation and, to a lesser extent, mechanical compaction. Hence, a significant portion of the original intergranular pore space (primary porosity) is usually preserved. The situation is quite different in carbonates. By and large, most of the primary porosity is obliterated during diagenesis and new secondary porosity created by mechanisms such as dissolution or fracturing. The resulting pore structure of carbonates is thus often much more complex than that of

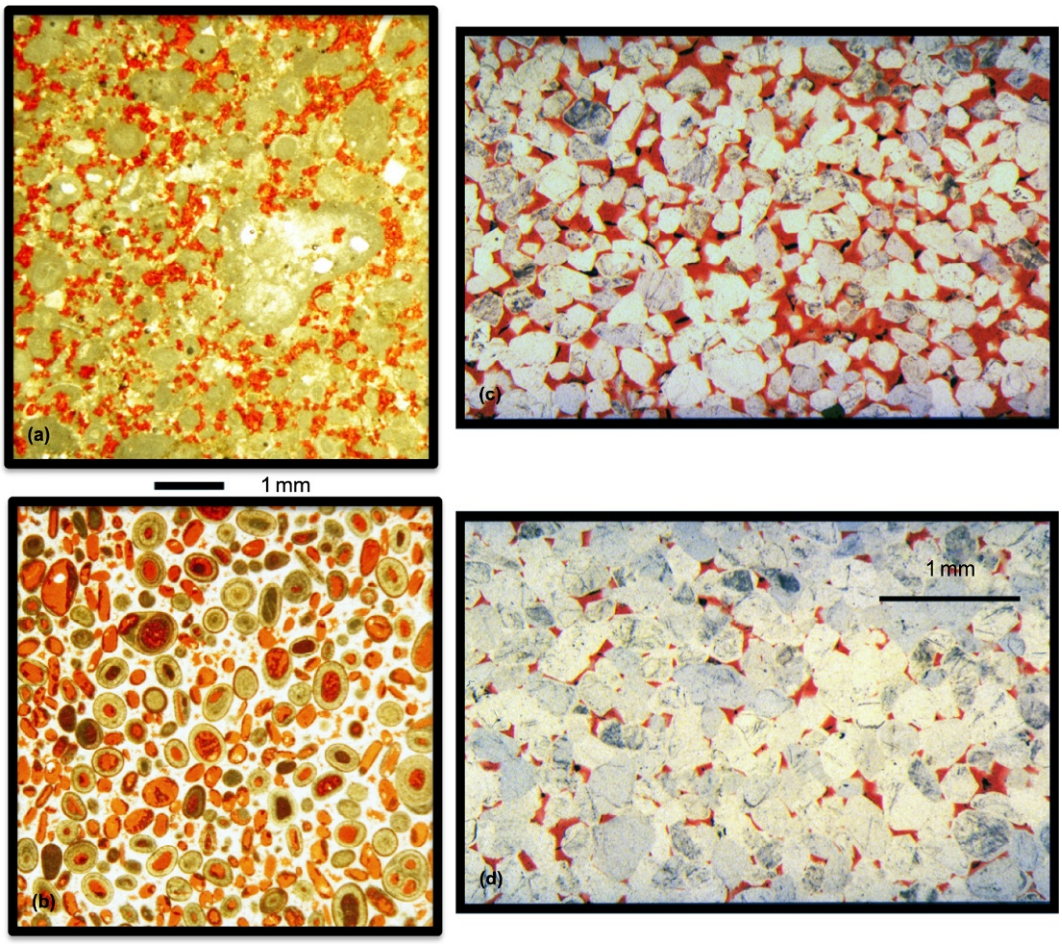


Figure 1 Transmitted light, optical microscope images (courtesy of M. Zamora) of two oolitic limestones, (a) porosity $\phi = 0.13$ and permeability $k = 6$ mD (i.e., 6×10^{-15} m^2) and (b) $\phi = 0.29$ and $k = 10$ mD, and two varieties of Fontainebleau sandstone, (c) $\phi = 0.20$ and $k = 3400$ mD and (d) $\phi = 0.043$ and $k = 2$ mD. In these rock samples, the connected pore space was saturated with red-dyed epoxy.

sandstones. In particular, very large pore size variations are more prevalent in carbonates than sandstones, although both types of rocks may contain sizeable amount of microporosity (often associated with micritic calcite matrix in carbonates and authigenic clay cement in sandstones). As an illustration, Figure 1 shows two oolitic limestones and two varieties of Fontainebleau sandstone. Despite the fact that they started with presumably similar oolitic sediments, the two limestones evolved completely differently during diagenesis. In one (Figure 1(a)), the oolites were totally cemented internally and all the porosity was located in the interoolite space, while in the other (Figure 1(b)), the opposite happened – the internal oolitic porosity was preserved and the interoolite space became almost entirely cemented. Fontainebleau sandstone is an unusually simple rock. It is composed of nearly 100% quartz and exists in many varieties covering an extremely wide range of porosities. In the examples shown in Figure 1(c) and 1(d), a decrease of porosity by a factor of 5 was observed corresponding to a drop in permeability by three orders of magnitudes (M. Zamora, personal communication).

We should also briefly mention the other important types of rocks present in the crust: igneous and metamorphic rocks such as granites, gabbros, schists, gneisses, or marbles (Bucher and Grapes, 2011). Since these rocks were subjected to high temperatures and high pressures during their formation, they tend to have very little porosity, often in the form of fissures. Thin microcracks are highly compliant and tend to increase the pressure dependence of the physical properties of rocks. Some of these rocks were intensely strained and developed highly anisotropic fabrics (Passchier and Trouw, 2005; Ramsay, 1967). In many rock formations, fluid flow at the macroscopic scale is controlled by fractures (e.g., Brace, 1981, see also the review by Guéguen et al., 1996). Fractures have an immense importance in many field applications. They require specific treatments and methods that are outside the scope of this chapter, and we refer the reader to the reviews of the vast literature on fractures and faults (e.g., Adler and Thovert, 1999; Berkowitz, 2002; Gudmundsson, 2011; Sahimi, 1995).

11.02.3 Microstructure

11.02.3.1 Geometry of the Pore Space

In general, the geomaterials considered in this volume come out as intricate arrangements of mineral grains, patches of

cement, and connected or disconnected pore space. Years of rock physics investigations have made it very clear that the microstructure of such materials (i.e., the geometry of the various phases) is a factor as important in controlling the medium properties as the individual properties of its constituents. The determination and characterization of the microstructure is therefore an extremely important task in petrophysics. In the past, microstructural parameters were measured on images of 2D rock sections and interpreted in terms of 3D quantities using the principles and methods of quantitative stereology (e.g., Underwood, 1970; see also Blair et al., 1996, for the use of two-point correlation functions). Quantitative stereology relies on statistics and combinations of simple measurements such as point counting or intercept length measuring. For example, the volume fractions of the various constituents are equal to the corresponding areal fractions measured on a section, provided the section is statistically representative (if it is not, additional sections must be used). To measure areal fractions, one can simply superpose a regular array or a random set of points on the micrograph and count the number of points falling in each phase. The precision of the measurement increases as the square root of the total number of points used. Similarly, the 3D surface-to-volume ratio of each phase can be inferred from the 2D perimeter-to-surface ratio, a quantity that is itself obtained from the number of phase bodies intersected by a set of test lines, per unit test line length ($S_V = 2P_L$, using Underwood's notations). When applied to the pore space, point counting yields the porosity ϕ, whereas the hydraulic radius r_h can be obtained from intercept length measurements (in this chapter, we define r_h as twice the inverse of the surface-to-volume ratio of the pore space, so that the hydraulic radius of a cylindrical pipe is equal to its radius; the factor 2 is often omitted in other studies). Notice that ϕ and r_h can also be measured using physical methods (as opposed to microstructure analysis methods) and, as will be discussed later, are widely used in models of permeability. Particulate systems (e.g., the solid part of a granular aggregate) can also be characterized using 2D quantitative stereology, although additional assumptions concerning the shape or size of the particles are often necessary (see Underwood, 1970, for more details). For example, finding the size distribution of packed ellipsoids requires the assumption that all the particles belonging to a given size class have the same shape (e.g., prolate or oblate ellipsoids with identical aspect ratios).

Only a limited number of geometric characteristics can be estimated using 2D quantitative stereology. Moreover, the microstructural quantities accessible may, in fact, not be the most relevant to the physical properties of the materials considered. In an attempt to evade the necessity to characterize a large number of possibly irrelevant pore geometry parameters, petrophysicists have proposed to construct virtual 3D pore structures such that random slices through them would be statistically identical to the sections actually observed and to numerically simulate fluid flow or electrical conduction through these digital porous media (e.g., Adler, 1992; Adler et al., 1990; Bentz and Martys, 1994; Biswal et al., 1999; Øren and Bakke, 2002; Sallès et al., 1993); the method was extended to the elastic properties (Poutet et al., 1996) and to multiscale structures (Biswal et al., 2009). The crucial question, of course, is to define what 'statistically equivalent' means. In this type of work, the spatial distribution of the phases is usually specified by a stochastic phase function $Z(\mathbf{x})$ defined as follows:

$$Z(\mathbf{x}) = 1 \text{ if the position vector x ends in the pore space, } Z(\mathbf{x}) = 0 \text{ otherwise} \quad [1]$$

(note that, here and in the rest of this chapter, bold characters will denote vectors or tensors). The statistical properties of Z are then assumed to be entirely described by its mean, that is, the porosity $\phi = \langle Z(\mathbf{x}) \rangle$, and the two-point correlation function $G(\mathbf{u})$:

$$G(\mathbf{u}) = \langle (Z(\mathbf{x}) - \phi)(Z(\mathbf{x}+\mathbf{u}) - \phi) \rangle / \phi(1-\phi) \quad [2]$$

where the brackets denote spatial averaging and $\mathbf{u}$ is the separation vector (in general, the field Z is fully correlated at $\mathbf{u} = 0$ and becomes more and more uncorrelated as $\mathbf{u}$ increases). However, it was found that media generated by matching ϕ and G do not always have the correct transport properties. According to Øren and Bakke (2002), the input of ϕ and G alone does not ensure proper simulation of the connectivity of the pore space. Thus, the numerical reconstruction method cannot be purely statistical but must include additional steps, for example, mimicking some of the physical processes thought to have actually occurred during the genesis of the rock considered (e.g., deposition and cementation; Øren and Bakke, 2002).

11.02.3.2 Topology

In addition to their dependence on metric microstructural characteristics such as pore size or pore length, the transport properties of porous rocks are strongly affected by the connectivity of the pore space (e.g., Adler, 1992; Dullien, 1992; Guéguen and Palciauskas, 1994; Sahimi, 1995). Although the concept of connectivity is intuitively clear, pore connectivity is not easy to define quantitatively. At the deepest level, connectivity is related to the topological invariants characterizing the pore space, and it has been proposed to use the first Betti number (or genus) as a measure of pore connectivity (DeHoff et al., 1972; Doughty and Tomutsa, 1997; Levitz, 2007; MacDonald et al., 1986; Scott et al., 1988; Vogel, 1997, 2002). The genus ζ of a multiply connected body is equal to the maximum number of cuts that can be made through it without separating it into two disconnected parts (e.g., the genus of a sphere is equal to zero, while a torus has a genus of one). However, ζ is difficult to use, in particular because it is an extensive property, increasing with the amount of material under consideration. In practice, it is convenient to measure the genus ζ as

$$\zeta = N_b - N_n + 1 \quad [3]$$

where N_b and N_n are the number of branches and nodes of the retract graph (often called the medial axis) obtained by thinning the pore space into a topologically equivalent skeleton. One great difficulty in practice is that the retract graph representing a particular multiply connected body is not unique. As an illustration, **Figure 2** shows three different graph elements that contribute the same $N_b - N_n$ to the graph genus and are therefore equivalent topologically. Furthermore, any branch can be divided into two segments separated by a new node

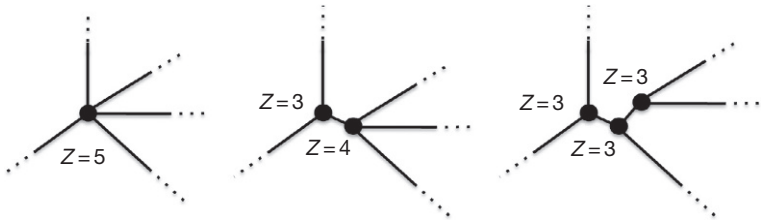


Figure 2 Three examples of topologically equivalent graph elements, equally contributing to the genus of the overall graph but having different mean coordination numbers. Algorithms used to determine retract graphs from 3D images of the pore space tend to produce configurations such as the right-end one, yielding values of the coordination number strongly biased toward 3. Rules for merging nodes are often used (i.e., moving toward the left in the figure), thus generating higher values of z. However, as far as we know, these rules are usually purely geometric and have not been justified on a physical basis.

and an arbitrary number of branches ending on a single node (i.e., dead-end branches) can be added to any node without changing ζ. Topologically equivalent retract graphs can therefore have different coordination numbers (i.e., the mean number of branches per node). The nonunicity of the retract graphs is a major source of ambiguity and uncertainty when retract graphs are used to identify individual pores and channels and measure their geometric characteristics. Another particularly difficult problem is that of characterizing the topology of materials with embedded multiscale pore structures, such as sandstones having part of the intergranular pore space filled up with microporous clay cement.

Although it is not a topological constant, the mean pore coordination number z has been proposed as an alternative way to quantify pore connectivity (e.g., Bernabé et al., 2010, 2011; Dullien, 1992). Different methods for calculating the retract graph from 3D images of the pore space have been proposed (e.g., Al-Kharusi and Blunt, 2007; Al-Raoush and Willson, 2005; Arns et al., 2004b; Baldwin et al., 1996; Jiang et al., 2007; Lindquist et al., 1996; Petford et al., 1999; Silin et al., 2004; Sok et al., 2002; Spanne et al., 1994; Vogel and Roth, 2001; Zhu et al., 2011). In its simplest form, the thinning algorithm used for constructing the retract graph leads to graphs containing only nodes with local coordination numbers equal to 3 and 1 (triple junctions and dead-end branches, respectively; e.g., Zhu et al., 2011). In the studies mentioned earlier, groups of nodes were sometimes merged together if the branches connecting them were considered too short (see **Figure 2**, going from right to left) or according to some other geometric criterion. Very different merging criteria have been used, yielding wide discrepancies in the estimated mean value of z. Establishing the most appropriate node merging criterion based on hydrodynamics arguments (as opposed to purely geometric ones) remains an open problem. One possible approach for this task is linked to the basic principle underlying the network simulation method, namely, that energy dissipation is assumed to occur exclusively in the branches (not in the nodes, which are only used to express the relevant conservation law). Accordingly, the two nodes attached to a particular branch can be merged without overly affecting the result of the network simulation, if the energy dissipated in the branch is negligibly small (e.g., this branch may be very thick and short and oppose negligible resistance to flow).

The topology of the solid phase in porous rocks may also be important in relation to other rock properties. For example, the way stress is transmitted from grain to grain in a granular aggregate strongly depends on the topology of the solid phase. If the grains are convex (or concavo-convex with limited concavity), the Delaunay tessellation (or its dual, the Voronoi tessellation) can be used to identify the force transmitting the grain network as well as the intergranular pore network (e.g., Bernal and Finney, 1967; Luchnikov et al., 1999; Medvedev et al., 2006). The relationship between the dual grain and pore networks is illustrated in **Figure 3**.

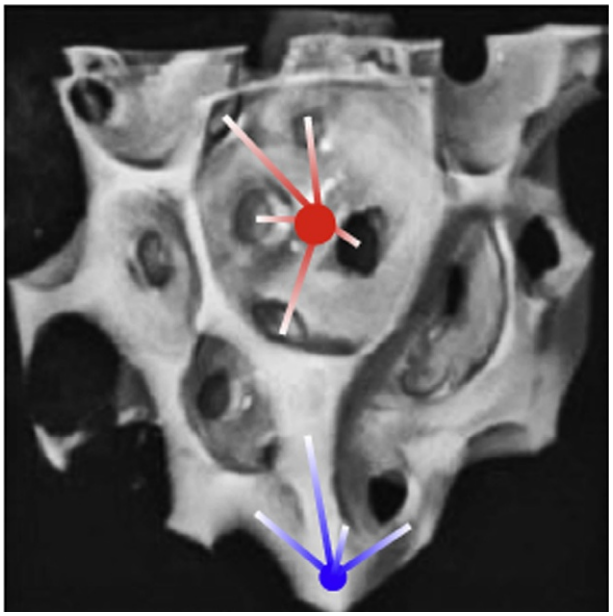


Figure 3 A 3D reconstruction of the intergranular pore space of a random packing of monodisperse spheres (courtesy Aste and Weaire, 2008). Pores are shown as an opaque body in gray color, while the spheres are 100% transparent. Small circular holes corresponding to stress-transmitting grain contacts are visible. The retract graphs of the pore and solid phases can be obtained from this type of 3D images as schematically illustrated in blue and red colors, respectively.

11.02.3.3 Summary

Recent advances in high-resolution 3D imaging are now allowing 3D geometric and topological characterization of the pore space. In particular, the pore–solid interface can be determined in many porous materials and directly incorporated in numerical simulations of physical processes such as fluid flow or

electrical conduction. Nevertheless, a number of open questions remain. For example, materials with multiscale, embedded pore structures (e.g., microporosity) are still difficult to characterize and model. The role of topology in controlling the physical properties of rocks is still poorly understood. The relation between topology and connectivity is yet to be clarified.

11.02.4 Single-Phase Flow

11.02.4.1 Darcy's Law, Permeability

Empirical evidence shows that steady-state, laminar flow of an interstitial fluid through a rigid, isotropic porous body is characterized by a linear relationship between the fluid volumetric flux $\mathbf{q}_f$ (i.e., volume per unit cross-section area and unit time) and the pressure gradient externally applied ∇p_f:

$$\mathbf{q}_f = -\frac{k}{\eta}\nabla p_f \qquad [4]$$

where k is the permeability and η the fluid viscosity. Equation [4] shows the isotropic form of Darcy's law, in which permeability appears as a scalar quantity. Its full anisotropic, tensorial form can be derived using the volume-averaging method (e.g., Gray and O'Neil, 1976; Neuman, 1977; Whitaker, 1986a), a brief description of which is given later.

Let us consider a representative elementary volume (REV) of statistically homogeneous, possibly anisotropic, and porous material (note that the method can also be applied to an inhomogeneous porous body comprising many REVs, provided that the properties of the medium change only gradually over the length of the body). The REV has a bulk volume V, contains a volume of fluid V_f, and is characterized by a length scale L assumed to be much greater than the characteristic pore-scale l:

$$l \ll L \qquad [5]$$

Assuming fluid incompressibility, the fluid velocity $\mathbf{u_f}$ must verify the continuity equation

$$\nabla\cdot\mathbf{u_f} = 0 \qquad [6]$$

and Stokes equation of motion

$$\eta\nabla^2\mathbf{u_f} = \nabla p_f \qquad [7]$$

where p_f denotes the local fluid pressure. In addition to the external boundary condition ($\mathbf{u_f}$ or p_f specified on the boundaries of the REV), the fluid must also obey the internal no-slip boundary condition (i.e., $\mathbf{u_f}=0$ on the solid–fluid interface A_f).

Volume averaging is then applied to eqns [6] and [7] and the boundary conditions. Two kinds of averages are used, namely, the bulk volume average,

$$\langle\psi\rangle = \frac{1}{V}\int_{V_f}\psi\,dV \qquad [8]$$

and the fluid volume average,

$$\langle\psi\rangle_f = \frac{1}{V_f}\int_{V_f}\psi\,dV = \frac{\langle\psi\rangle}{\phi} \qquad [9]$$

where ψ is a fluid-defined quantity, ϕ is the porosity (i.e., $\phi = V_f/V$), the brackets denote averaging, and the subscript f refers to the fluid phase. The volume-averaging theorem states (e.g., Slattery, 1967)

$$\langle\nabla\psi\rangle = \nabla\langle\psi\rangle + \frac{1}{V}\int_{A_f}\mathbf{n}\psi\,dA \qquad [10]$$

where A_f is the solid–fluid interface and $\mathbf{n}$ is the unit vector, normal to the interface in the outward direction from the fluid phase. Under volume averaging, eqns [6] and [7] become

$$\nabla\cdot\langle\mathbf{u_f}\rangle_f = 0 \qquad [11]$$

and

$$\nabla\langle p_f\rangle_f = \frac{\eta}{V_f}\int_{A_f}\mathbf{n}\cdot\nabla\delta\mathbf{u_f}dA - \frac{1}{V_f}\int_{A_f}\mathbf{n}\delta p_f dA \qquad [12]$$

where simplifications (mostly based on inequality 5) were made and $\mathbf{u_f}$ and p_f were decomposed into $\mathbf{u_f}=\langle\mathbf{u_f}\rangle+\delta\mathbf{u_f}$ and $p_f=\langle p_f\rangle+\delta p_f$; Whitaker (1986a) showed that closure of eqn [12] can be carried out by relating the local deviations $\delta\mathbf{u_f}$ and δp_f to $\langle\mathbf{u_f}\rangle_f$ by means of the following linear mapping:

$$\delta\mathbf{u_f} = \mathbf{B}\cdot\langle\mathbf{u_f}\rangle_f \quad \text{and} \quad \delta p_f = \eta\mathbf{b}\cdot\langle\mathbf{u_f}\rangle_f \qquad [13]$$

where tensor $\mathbf{B}$ and vector $\mathbf{b}$ satisfy the following boundary value problem:

$$\nabla\cdot\mathbf{B} = 0 \qquad [14]$$

$$-\nabla\mathbf{b} + \nabla^2\mathbf{B} = \frac{1}{V_f}\int_{V_f}(-\nabla\mathbf{b} + \nabla^2\mathbf{B})dV \qquad [15]$$

$$\langle\mathbf{b}\rangle_f = \langle\mathbf{B}\rangle_f = 0 \qquad [16]$$

The internal boundary condition reduces to $\mathbf{B}+\mathbf{I}=0$ on A_f (where $\mathbf{I}$ is the unit tensor) and the original external boundary conditions must be volume-averaged and transformed in a way consistent with eqn [13]. In the end, an equation in the form of Darcy's law is obtained:

$$\langle\mathbf{u_f}\rangle = -\frac{\mathbf{K}}{\eta}\cdot\nabla\langle p_f\rangle_f \qquad [17]$$

with the permeability tensor given by

$$\mathbf{K}^{-1} = \frac{-1}{\phi V_f}\int_{A_f}\mathbf{n}\cdot(\nabla\mathbf{B} - \mathbf{I}\mathbf{b})dA \qquad [18]$$

$\mathbf{K}$ was shown to be a symmetric tensor and a material property (i.e., independent on the external boundary conditions) only if the boundary value problem considered is periodic (Whitaker, 1986a). However, it can be argued that an arbitrary problem on the original REV can always be made part of a larger periodic system, thus preserving the properties of $\mathbf{K}$. Neuman (1977) gave another derivation of the properties of $\mathbf{K}$ using superpositions of solutions of eqns [6] and [7] (in other words, relying on the linearity of Stokes equation).

Thus, at the deepest level, the derivation sketched earlier depends on the linearity of the fluid equation of motion and on separation of scales (i.e., inequality 5). The other assumptions, such as statistical homogeneity or steady-state conditions, are less important and can be relaxed to some extent. As already stated, the method remains valid for materials with gradually varying properties. The volume-averaging method can also be applied to time-dependent flow as long as the nonlinear convective terms in the Navier–Stokes equation are negligible. Fluid storage and the mechanical coupling between

the fluid and the solid matrix play an essential role in unsteady-state flow conditions and cannot be neglected as was done for steady-state flow. To analyze time-dependent flow, a very appealing idea is to take the point of view of spectral analysis and treat permeability as a frequency-dependent property (e.g., Auriault et al., 1985; Johnson, 1989; Johnson et al., 1987). Two regimes can be identified: (i) the low-frequency, viscous regime, in which $k(\omega)$ asymptotically approaches the steady-state permeability k, and (ii) the high-frequency, inertial regime, in which $k(\omega)$ decreases as $1/\omega$ and becomes $\pi/2$ out of phase with the applied harmonic pressure gradient (i.e., if $k(\omega)$ is expressed as a complex number, its real part vanishes with increasing ω faster than its imaginary part). The transition occurs at a critical frequency ω_c approximately given by

$$\omega_c = \frac{8\eta}{l^2\rho_f} \qquad [19]$$

where ρ_f is the fluid density and l refers to a characteristic pore length scale, which might be identified with the geometric or the dynamic hydraulic radius, r_h and Λ, respectively (for a precise definition of Λ, see Johnson et al., 1987). Assuming that hydraulic and electrical tortuosities are identical, eqn [19] can be written in an alternative form, $\omega_c = \eta/(\rho_f F k)$, where F is the electrical formation factor. Finally, note that the dynamic permeability $k(\omega)$ is closely related to the Biot slow compression wave (e.g., Johnson et al., 1982; see also Bernabé, 2009, who modeled the propagation of Biot slow waves in heterogeneous networks of pipes).

Fluid flow through porous media, as expressed by Darcy's law, is a dissipative process, in which energy dissipation is exclusively caused by the viscous forces acting inside the fluid. However, other dissipation mechanisms may play a role in some situations. For example, transport of ionic charges owing to electrokinetic coupling introduces Joule dissipation and, thus, impedes fluid flow without changing the geometry of the pore space (Pride, 1994; Revil and Glover, 1997; Revil and Pessel, 2002; Revil and Linde, 2006; see also Bernabé, 1998, who simulated streaming potential in heterogeneous networks of pipes).

11.02.4.2 Measurement of Permeability

Although we cannot undertake a detailed, comprehensive discussion of the laboratory and field techniques for measuring permeability, a few words are necessary. In the laboratory, uniaxial fluid flow is usually generated through fully saturated, cylindrical samples inserted in an impermeable jacket and subjected to controlled conditions of confining and pore fluid pressure. In order to ensure representativeness of the measurement, the sample must be statistically homogeneous (however, geologic formations may contain fractures and/or self-similar heterogeneities, so that even a large number of laboratory measurements may not be sufficient to fully describe them; e.g., Boutt et al., 2012; Guéguen et al., 1996). Violations of the sample homogeneity condition will affect a permeability measurement in different ways depending on the flow technique used. The most common technique is to apply a constant fluid pressure difference between the end faces of the sample and, therefore, generate steady-state flow. Under these conditions, an apparent value of the scalar permeability is directly obtained from the measured flow rate by means of the scalar form of Darcy's law (eqn [4]). If the homogeneity condition is not satisfied, the apparent permeability measured is not representative of the medium, but unfortunately, this failure cannot be diagnosed from the measurement itself.

Steady-state measurements are usually time-consuming, and faster unsteady-state flow techniques, such as the pulse decay technique (e.g., Boitnott, 1997; Brace et al., 1968), have been proposed in the past. Other methods, each with its own advantages and drawbacks, include the harmonic flow, constant injection rate, and the linear pressurization methods (Fischer, 1992; Song and Renner, 2006; Song et al., 2004, respectively). The unsteady-state flow response to an arbitrary input stimulation (whether a pressure pulse, a pressure harmonic oscillation, or something else) involves not only the material permeability but also its storage capacity (e.g., Fischer, 1992). For accurate measurements of the material storage capacity, the storativity of the measuring system must be comparable to or smaller than the sample storativity (e.g., Fischer, 1992; Song and Renner, 2006; Song et al., 2004). In most of the unsteady-state techniques, the determination of permeability and storage capacity is indirectly done by comparison of the recorded flow response to the solution of the corresponding 1D flow problem, assuming a homogeneous porous medium. As a consequence, unsteady-state techniques are, by and large, capable of identifying violations of the homogeneity condition mentioned earlier. Indeed, uninterpretable, pathological flow responses are generally produced in overly inhomogeneous samples (see the striking examples reported by Kamath et al., 1992).

Another crucial assumption implied by using a boundary condition of no lateral flow in laboratory measurements (i.e., impermeable jacket) is that the material is isotropic and satisfies the scalar form of Darcy's law. Severe interpretation problems are to be expected when running standard laboratory measurements in strongly anisotropic samples, with the principal axes obliquely oriented with respect to the nominal flow direction (e.g., Renard et al., 2001). Laboratory measurements of the full permeability tensor require multiple steady-state flow experiments in different directions and a specific interpretation technique based on numerical flow simulations (Renard et al., 2001). Note also the interesting study of Clavaud et al. (2008), who inferred the permeability tensor and its relationship with the material pore structure, from images of a tracer front propagating through the samples during steady-state flow experiments.

Permeability can also be measured in the field using borehole forced-flow methods (e.g., Butler, 2005). The main differences with the laboratory methods mentioned earlier are (i) the unbounded nature of the aquifers or reservoirs into which flow is induced and (ii) the approximately radial geometry of the flow field near a borehole source. Consequently, the field methods involve time-dependent flow and are very sensitive to permeability heterogeneities. The spatial distribution of permeability can also be inferred from tracer tests in forced or natural flow conditions as well as from surface or borehole geophysical measurements, which rely on established relationships of permeability to other rock properties (see the fairly complete monograph edited by Rubin and Hubbard, 2005). For example, self-potential measurements can be interpreted in some cases in terms of underground water flow and allow

determination of the permeability structure (e.g., pumping test (Rizzo et al., 2004), natural flow (Revil et al., 2005), and periodic forced flow (Maineult et al., 2008)).

11.02.4.3 Permeability Models

It is not possible in the space allowed for this chapter to review all the permeability models that have been proposed in the past (e.g., see Guéguen et al., 1996, for a short review). Instead, we will concentrate on two: (i) the equivalent channel model independently devised by Paterson (1983) and Walsh and Brace (1984) and hereafter referred to as PWB and (ii) our own recent attempt to take pore-scale heterogeneity and pore connectivity into account (Bernabé et al., 2010, 2011).

Although it could, in principle, be extended to anisotropic media, the PWB model considers an isotropic porous material and assumes that, as far as the transport properties are concerned, its pore space at the microscopic scale can be replaced by a single equivalent channel as schematically shown in **Figure 4** (inspired by Figure 1 in Walsh and Brace, 1984). The equivalent channel should, of course, be representative of the pore geometry in addition to permeability, electrical resistivity, and any other transport properties considered. Hence, the ratio of the cross-sectional surface area of the equivalent channel to that of the lump of porous material in which it is embedded (i.e., A^*/A in **Figure 4**) must be equal to the porosity ϕ. Likewise, the equivalent channel must have the same pore volume-to-surface ratio (i.e., hydraulic radius) as the porous material that it is meant to represent. A third geometric constraint has to be included: the tortuosity τ, the ratio of the length of the streamlines to the length of the porous body in the nominal flow direction, must be preserved in the equivalent channel description (i.e., $l/x = l^*/x^* = \tau$ in **Figure 4**). Under these premises, it is not difficult to see that permeability k and resistivity formation factor F are given by

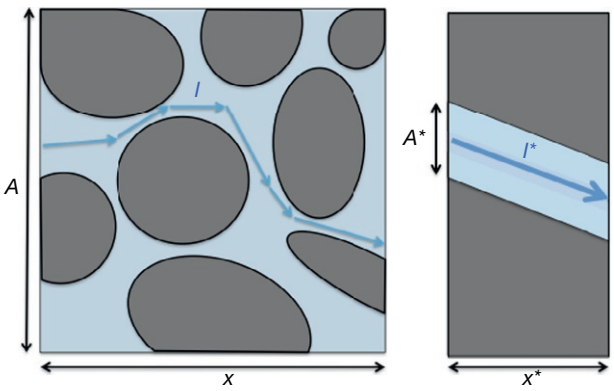


Figure 4 A schematic representation of the equivalent channel (based on Figure 1 of Walsh and Brace, 1984). The left picture shows an idealized 2D image of a porous body of length x and cross-section surface area A (perpendicularly to the picture plane). A tortuous flow line of length l is included in the pore space. The picture to the right depicts the equivalent channel. It has a length l^* corresponding to a nominal length x^* and a cross-section surface area equal to A^* (with $A^*/A = \phi$).

$$k = \frac{Cr_h^2\phi}{\tau^2} \quad [20]$$

and

$$F = \frac{\tau^2}{\phi} \quad [21]$$

where C is a shape factor (equal to 1/8 for a straight, cylindrical pipe and 1/12 for a flat, crack-like slit). Since, as will be discussed later, tortuosity is a rather ill-defined parameter, it is practical to combine eqns [20] and [21] to obtain

$$k = \frac{Cr_h^2}{F} \quad [22]$$

Notice that, in eqns [21] and [22], F is classically defined as the ratio of the porous material resistivity to that of the fluid saturating it, under the strict assumption that surface conduction effects either are absent or have been eliminated using a suitable correction (the same assumption will apply to the Bernabé et al.'s model described in subsequent paragraphs; for a more general definition of F, see Revil and Glover, 1997, 1998; Revil and Leroy, 2004). Paterson (1983) tested eqn [22] (using the median value $C = 1/10$) for a set of materials, for which experimental data on r_h, ϕ, F, and k were available. He found that k was systematically overestimated by a factor of about 2 in the case of granular materials (with mid to high porosities) and by one or two orders of magnitude for low-porosity, fissured granites. These results are fairly typical of the PWB model. They are considered satisfactory in the case of the granular materials since it is usually difficult to predict permeability to significantly better than a factor of 2. This relative success, in addition to the extreme simplicity of the model, explains the broad popularity of PWB (and of all similar models, often referred to as Kozeny–Carman models). However, the systematic overestimation noted earlier and the poor fit obtained for tight, fissured rocks point to some fundamental problems.

The shape factor C can be directly ruled out as a plausible source of error since it cannot spread outside the narrow range mentioned earlier (i.e., 1/8 to 1/12). The tortuosity factor is a different matter (see Dullien, 1992 and the excellent review by Clennell, 1997). Clearly, the ratio l/x cannot be constant for all possible streamlines. The streamlines near the (possibly rough) pore walls must be much more tortuous than the ones running at the center of the pore channels. It may be argued that the wall streamlines (or, more properly, stream tubes) carry much less fluid than the central ones and must therefore be given less weight. Hence, a weighted averaging procedure is implied in the definition of τ, but, as far as we know, identifying it has never been attempted. Accurately estimating the tortuosity factor directly from 3D pore geometry images is thus unfeasible at present and may remain so even when the streamlines are determined through flow simulations. It is not surprising then, if the PWB model is often reduced to eqn [22]. But, even in this form, problems arise. In particular, eqn [22] implicitly assumes that the (averaged) tortuosity of hydraulic streamlines and electrical current lines are identical. This is obviously true for the straight, constant cross-section equivalent channel (see **Figure 4**), but it is far from assured in more complex porous media since the no-slip internal boundary condition applies to

fluid flow but not to electrical conduction. Nevertheless, inferring the tortuosity factor from electrical measurements is commonly done and it must be recognized that it does not usually lead to gross contradictions (e.g., Johnson et al., 1987).

Most importantly, PWB ignores the pore-scale variability of the materials. An unspecified averaging procedure is implied in the model, without any guarantee as to its validity. Another problem is the oversimplified treatment of pore connectivity. Strictly speaking, the equivalent channel is representative of the hydraulically conducting pore space only. According to PWB, it is sufficient to simply exclude isolated voids and dead-end pores, assuming they can be unambiguously identified, when estimating ϕ and r_h (e.g., Paterson, 1983). The PWB approach neglects the effect of topology, that is to say, the fact that porous media having identical porosities, pore dimensions, and so on, but differing in topology, have different permeabilities.

In two recent studies, Bernabé et al. (2010, 2011) tried to model the effects of pore-scale variability and pore connectivity. Their approach was to numerically simulate fluid flow and electrical conduction through idealized porous media with controlled pore-scale variability and pore connectivity, identify robust relationships between the input pore parameters and the transport properties, and generalize the previously established relationships to real porous media. For this purpose, heterogeneous, variably connected pipe networks were constructed by randomly populating the networks with pipes according to a preset occupancy probability p and randomly assigning the values of the pipe radii r_i according to a preselected statistical distribution. Parameterization was achieved by using the normalized pipe radius standard deviation $\sigma_r/\langle r_i\rangle$ and the coordination number z (i.e., mean number of pipes per node) as measures of pore-scale variability and pore connectivity, respectively. Following a standard percolation theory approach, the network realizations were constructed using different underlying lattices (i.e., simple cubic, SC; body-centered cubic, BCC; and face-centered cubic, FCC, with maximum z of 6, 8, and 12, respectively). The robustness of any relationship identified was assessed by its independence on lattice type (this is usually called 'universality' in percolation theory jargon). The coordination number z and occupancy probability p were linearly related for each lattice type, implying that the percolation theory analyses usually performed in terms of p can equivalently be carried out using z. Moreover, z at the percolation threshold has a single critical value $z_c = 1.5$ for all lattices considered (see also Dullien, 1992, or Sahimi, 1995). Hence, z_c is a 'universal' quantity, unlike the critical occupancy probability p_c, which varies with lattice type (i.e., $p_c = 0.249$, 0.180, and 0.120 for SC, BCC, and FCC, respectively).

The main result of the network simulations was that, for a given value of the pore-scale variability measure $\sigma_r/\langle r_i\rangle$, k and $1/F$ were well fitted by 'universal' power law functions (i.e., $k \propto (z-z_c)^{\beta}$ and $1/F \propto (z-z_c)^{\gamma}$ with exponents and prefactors independent on lattice type; see examples in **Figure 5**). These power laws are similar to the critical power laws of $(p-p_c)$ identified in various applications of percolation theory (e.g., Hunt and Ewing, 2009; Sahimi, 1995). They have the great advantage, however, of being valid for all values of z (except maybe in a close vicinity of z_c), whereas the critical power laws of percolation theory hold only near the percolation threshold

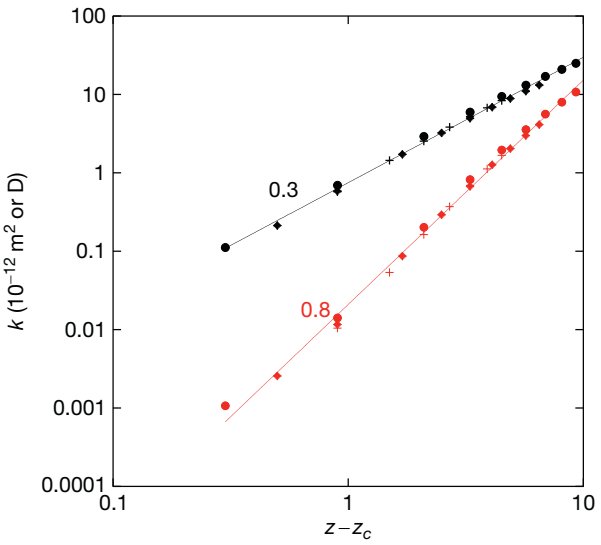


Figure 5 Two examples of the power laws $k \propto (z-z_c)^{\beta}$ numerically determined in Bernabé et al. (2010, 2011). Two clusters of data points are shown, corresponding to a low and high level of pore-scale variability (i.e., $\sigma_r/\langle r_i\rangle = 0.3$ and 0.8, black and red symbols, respectively). In each cluster, the data points were obtained using different types of lattices (FCC, BCC, and SC networks are indicated by solid circles, diamonds, and crosses, respectively). The power laws are visibly well defined (straight lines with little scatter) and, most importantly, independent of lattice type, so they can be considered universal in the sense of percolation theory (see Bernabé et al., 2010, 2011, for more details).

(i.e., for $(p-p_c)$ approaching zero). The range of applicability is thus very large and not limited to materials at the onset of pore disconnection. The observed power laws lead to

$$k = C_k\left(\frac{r_h}{l}\right)^2 (z-z_c)^{\beta} r_h^2 \qquad [23]$$

and

$$\frac{1}{F} = C_F\left(\frac{r_h}{l}\right)^2 (z-z_c)^{\gamma} \qquad [24]$$

where l denotes the pipe length and the prefactors and exponents (i.e., C_k, C_F, β, and γ) are functions of $\sigma_r/\langle r_i\rangle$ that have to be empirically determined from the results of the network simulations. We emphasize that, unlike the tortuosity factor τ, the coordination number z characterizes the pore space geometry and, therefore, applies equally to fluid flow and electrical conduction. Hence, no objections can be raised to combining eqns [23] and [24] into

$$k = C\left(\frac{r_h}{l}\right)^{2(1-\alpha)}\left(\frac{1}{F}\right)^{\alpha} r_h^2 \qquad [25]$$

where $C = C_k/C_F^{\alpha}$ and $\alpha = \beta/\gamma$. Notice that eqn [25] reduces to eqn [22] when the exponent α is equal to unity, which occurs for $\sigma_r/\langle r_i\rangle = 0$ or, in other words, for materials in which the pores are all the same (for perfectly homogeneous networks of cylindrical pipes, C is equal to 1/8, a direct consequence of Poiseuille formula). The exponent α was observed to gradually increase with increasing pore-scale variability, up to 1.7 for extremely heterogeneous networks ($\sigma_r/\langle r_i\rangle = 1.05$), while C decreased to as low as 1/15.

Testing eqns [23]–[25] is difficult at the present time. Reports of measurements of ϕ, k, and F made on the same sample or set of samples can be easily found, but very few also provide microstructural data allowing estimation of r_h, l, and $\sigma_r/\langle r_i\rangle$. Measurements of z are extremely rare. In the end, the model could only be checked on unconsolidated granular media (well-sorted glass beads and sands and binary mixtures of sands) and one rock, Fontainebleau sandstone. The results were encouraging but by no means certain, owing to large uncertainties in the estimates of some of the parameters. The best fit was obtained for the glass beads because their simple, smooth, spherical shape allowed accurate determination of the microstructural parameters (in particular, the established value of the pore coordination number for well-sorted unconsolidated granular media, $z=6$, was assumed). The case of Fontainebleau sandstone is particularly interesting because varieties of this rock exist in a very wide range of porosities and variations in porosity correspond to changes not only in z, r_h, l, and $\sigma_r/\langle r_i\rangle$ but also in pore shape. The pores tend to flatten as porosity decreases and become crack-like at porosities lower than 10%. It turns out that the model fitted the experimental data satisfactorily only when the observed changes in pore aspect ratio were incorporated (for a more detailed discussion, see Bernabé et al., 2011).

11.02.4.4 Summary

The validity of Darcy's law is unanimously accepted and so is the definition of permeability as a material property and a sole function of pore geometry. However, finding out which aspects of pore geometry are most important for fluid flow is still debated today. A lot of attention has been traditionally devoted to porosity (i.e., the total amount of pore space) although its relationship to permeability cannot be all that strong. Indeed, nonconductive pores (even if connected as, for instance, dead-end pores) can be added or removed arbitrarily without affecting permeability. The fact that not all pore space is equally important for fluid flow was highlighted by Paterson (1983), who pointed out that isolated and dead-end pores had to be excluded from determinations of both porosity and hydraulic radius (but note that all physical measurements of porosity, using picnometry, e.g., do include dead-end pores automatically). Pore size (in the sense of pore aperture) is obviously a crucial quantity but may not be easy to characterize, especially in materials with very high pore-scale variability (e.g., carbonate rocks). Pore-scale variability is also affected by the extraordinarily complex pore shapes observed in microstructural images. For the purpose of modeling permeability, we can only assume, at this point, that pore shape characterization may be reducible to a few parameters such as cross-sectional aspect ratio. The problem of pore-scale variability is compounded by the fact that the geologic bodies of interest in field studies are macroscopically heterogeneous (e.g., Guéguen et al., 1996, and references therein). Finally, as also mentioned in **Section 11.02.3**, modeling of pore connectivity remains an open question. In particular, there is presently no consensus as to what the most appropriate measure of pore connectivity may be and how to incorporate it into a permeability model.

11.02.5 Solute Transport

11.02.5.1 General Considerations

Pollution of shallow aquifers is a great threat to the freshwater supply in many areas in the world. The transport of soluble contaminants by underground waters is therefore a topic of primary importance to hydrologists. Solute transport, often coupled to chemical interactions between rock and fluid, is also a major mass transfer mechanism in many geologic processes. Reactive transport is too complex, however, to include in this chapter. We will only consider passive transport and assume that fluid flow is unaffected by changes in solute concentration. In this case, solute transport at macroscopic scales is classically described as the combination of two mechanisms, convection and dispersion, expressed by the convection–dispersion equation (e.g., Dagan, 1989; Gelhar, 1993):

$$\frac{\partial C}{\partial t} + \nabla\cdot(\mathbf{U}C) = \nabla\cdot(\mathbf{D}\cdot\nabla C) \qquad [26]$$

where C is the tracer concentration, $\mathbf{U}$ the seepage velocity, and $\mathbf{D}$ the dispersion tensor (note that we used capital letters here to emphasize that the quantities involved are all macroscopic, averaged quantities). If $\mathbf{U}$ and $\mathbf{D}$ are constant in space, eqn [26] becomes

$$\frac{\partial C}{\partial t} + \mathbf{U}\cdot\nabla C = \mathbf{D}\cdot\nabla^2 C \qquad [27]$$

In the case of transport through an isotropic porous medium, the tensor $\mathbf{D}$ has two independent components corresponding to the directions parallel and perpendicular to the flow (for the anisotropic case, see Guin et al., 1972, and references therein).

Equation [27] predicts that the center of mass of a tracer plume travels at a constant velocity equal to $\mathbf{U}$ (convection) and that the plume spreads according to a Fickian process characterized by $\mathbf{D}$ (dispersion). Furthermore, the longitudinal and transversal dispersion coefficients, D_L and D_T (i.e., the components of $\mathbf{D}$), are assumed to be proportional to the seepage velocity, $D_L = A_L\,|\mathbf{U}|$ and $D_T = A_T\,|\mathbf{U}|$, where the longitudinal and transversal dispersivities, A_L and A_T, are considered material constants. This description is supported to first order by laboratory and field measurements (e.g., Bear, 1972; Gelhar, 1993, and references therein), but a number of observations have led to serious objections: (i) The comprehensive compilation of field measurements by Gelhar et al. (1992) demonstrated that A_L and A_T tend to increase with the duration and/or length scale of the measurements, and (ii) recorded time series of tracer concentration at observation points in laboratory experiments (i.e., breakthrough curves) displays significant discrepancies with $C(t)$ calculated using eqn [27] (e.g., Charlaix et al., 1987; Cortis and Berkowitz, 2004; Gouze et al., 2008; Levy and Berkowitz, 2003). The analysis of the observed breakthrough curves showed that the transport of a tracer pulse (initially injected at $t=0$) produces asymmetric breakthrough curves at the observation points, with long trailing tails, instead of the Gaussian curves predicted by eqn [27]. Alternative non-Fickian dispersion models have been proposed (e.g., Berkowitz and Scher, 1995).

Whether Fickian or not, dispersion is a very complex process, combining the effects of the fluctuations in local

convection velocity and of molecular diffusion (or pore-scale dispersion, i.e., an upscaled diffusive process such as Taylor dispersion). For example, Bear (1972) distinguished 'mechanical' and 'hydrodynamic' dispersion, the former referring to tracer spreading produced by the heterogeneity of the fluid velocity field, while the latter to pore-scale dispersion. The relative importance of each mechanism is expected to vary depending on the heterogeneity of the porous body considered, but it is usually believed that fluid velocity fluctuations, by far, have a predominant effect and lead to Fickian dispersion (at least, in the long time limit). Models relating the long-term dispersivities to the statistics of the permeability fluctuations have been published (e.g., Dagan, 1989; Gelhar, 1993; Naff, 1990; see also Burr et al., 1994, for a test of these models based on numerical simulations). Although molecular diffusion is a very slow transport mechanism, it allows solute particles to enter regions in the pore space with nearly zero fluid velocities. These stagnant regions act as tracer traps and are thought to produce non-Fickian breakthrough curves (e.g., Berkowitz and Scher, 1995; Cortis and Berkowitz, 2004; Gouze et al., 2008; Levy and Berkowitz, 2003). The occurrence of stagnant regions is strongly related to flow channeling and, therefore, to the heterogeneity of the porous body (Berkowitz and Scher, 1995).

11.02.5.2 Taylor Dispersion

The simplest example of the phenomena discussed earlier and the only one that can be completely solved analytically is Taylor dispersion, that is, the mixing process resulting from molecular diffusion and convection in a stationary, nonuniform, fluid velocity field. A detailed presentation of Taylor dispersion is helpful to illustrate the methods used for modeling solute transport through porous media (in particular, those based on the Fokker–Planck equation) and the problems encountered. Following the work of Van Den Broeck (1982, 1990), we consider an infinite tube of radius a, through which a steady-state Poiseuille flow is established:

$$u(r) = 2\langle u\rangle\left(1 - \frac{r^2}{a^2}\right) \qquad [28]$$

where x and r denote the axial and radial coordinates and the fluid velocity u is everywhere parallel to the x-axis. The motion of a solute particle located, at time t, at point $\{x(t), r(t)\}$ is the sum of passive convective motion at the local fluid velocity and of Brownian diffusive motion with a molecular diffusion coefficient D_m:

$$\frac{\partial x(t)}{\partial t} = u(r(t), t) + \psi_x(t) \qquad [29]$$

and

$$\frac{\partial r(t)}{\partial t} = \psi_r(t) \qquad [30]$$

where x and r now indicate the position of the particle and the stochastic functions ψ_x and ψ_r represent the diffusive white noise displacement in the axial and radial direction causing Brownian motion (note that, although ψ_x is Gaussian by definition, ψ_r is not because a reflecting boundary condition must be added at $r=a$ to prevent solute particles from leaking out of the tube). We must have

$$\langle\psi_x(t)\psi_x(t')\rangle = 2D_m\delta(t-t') \qquad [31]$$

and

$$\langle\psi_r(t)\psi_r(t')\rangle = 2D_m\delta(t-t') \qquad [32]$$

where δ denotes the Dirac function.

The main simplifying feature of the situation described earlier is that the radial motion is independent of the axial coordinate, a characteristic obviously not shared by solute transport through porous media. The statistics of the radial motion can be conveniently obtained using the Fokker–Planck representation of eqn [30]:

$$\frac{\partial P(r,t|r_0,t_0)}{\partial t} = D_m\nabla^2 P(r,t|r_0,t_0) \qquad [33]$$

with initial and boundary conditions $P(r,t_0|r_0,t_0)=\delta(r-r_0)$ and $\frac{\partial P(r,t|r_0,t_0)}{\partial r}=0$ at $r=a$. In eqn [33], $P(r,t|r_0,t_0)$ is the conditional probability of a particle position r at time t given r_0 at t_0. The solution of eqn [33] is known to be (Van Den Broeck, 1982)

$$P(r,t|r_0,t_0) = \frac{2}{a^2} + \frac{2}{a^2}\sum_{n=1}^{\infty}\exp\left[-\frac{D_m\mu_n^2(t-t_0)}{a^2}\right]\frac{J_0\left(\frac{\mu_n r}{a}\right)J_0\left(\frac{\mu_n r_0}{a}\right)}{\mu_n^2 J_0^2(\mu_n)} \qquad [34]$$

where μ_n are the positive roots of $J_1(\mu_n)=0$ (J_0 and J_1 are standard Bessel functions). We next compute the axial position $x(t)$ by integration of eqn [29]:

$$x(t) = x(0) + \int_0^t u(r(\tau),\tau)\mathrm{d}\tau + \int_0^t \psi_x(\tau)\mathrm{d}\tau \qquad [35]$$

We can then calculate the second moment $\langle\delta x^2(t)\rangle = \langle(x(t)-\langle x(t)\rangle)^2\rangle$ as

$$\langle\delta x^2(t)\rangle = 2D_m t + \int_0^t\int_0^t [\langle u(r(\tau),\tau)u(r(\tau'),\tau')\rangle - \langle u(r(\tau),\tau)\rangle\langle u(r(\tau'),\tau')\rangle]\mathrm{d}\tau\mathrm{d}\tau' \qquad [36]$$

where the averages can be estimated using the conditional probabilities of eqn [34] and the assumption that the initial radial distribution of solute particles is uniform and remains so at all times (Van Den Broeck, 1982):

$$\langle\delta x^2(t)\rangle = 2D_m t + 2\frac{a^2 t}{D_m}\sum_{n=1}^{\infty}\frac{64\langle u\rangle^2}{\mu_n^6} - 2\frac{a^4}{D_m^2}\sum_{n=1}^{\infty}\frac{64\langle u\rangle^2}{\mu_n^8}\left[1-\exp\left(-\frac{D_m\mu_n^2 t}{a^2}\right)\right] \qquad [37]$$

In the short time limit, eqn [37] predicts a transitory non-Fickian dispersion regime given by

$$\langle\delta x^2(t)\rangle \simeq 2D_m t + \frac{\langle u\rangle^2 t^2}{3} - \frac{8D_m\langle u\rangle^2 t^3}{3a^2} + \cdots \qquad [38]$$

The duration of the transient regime is on the order of $a^2/(16D_m)$, corresponding to the time needed for a solute particle to sample the entire cross-sectional fluid velocity field. We note

that the transient regime may persist a long time since D_m is small. Obviously, the occurrence of a long, non-Fickian, transient regime should also be expected for solute transport through porous media. In the long time limit, the longitudinal dispersion described by eqn [37] asymptotically approaches a Fickian regime, $\langle \delta x^2(t) \rangle \simeq 2D_{\text{Taylor}}t$, known as Taylor dispersion, with an effective dispersion coefficient D_{Taylor} equal to

$$D_{\text{Taylor}} = D_m + \frac{\langle u \rangle^2 a^2}{48 D_m} \quad [39]$$

The approach depicted earlier (i.e., eqns [29]–[36]) is, by no means, limited to the Poiseuille velocity field of eqn [28]. The only required condition is that the fluid velocity field must be coaxial and constant in the axial direction, ensuring a diffusive transversal motion decoupled from the convective axial motion (as in eqn [30]). Arbitrary fluid velocity fields obeying this condition will therefore yield expressions for the effective dispersion coefficient including a dependence on the square of the mean velocity as in eqn [39]. For example, this result is found for an infinite, layered porous body in the case of steady-state flow parallel to the layers (each assumed to be homogeneous).

When the decoupling condition mentioned earlier is not valid, the relation of the dispersion coefficient to the fluid velocity changes. For example, Bruderer and Bernabé (2001) simulated convection along the local streamlines and molecular diffusion (as in eqns [29]–[32]) in heterogeneous networks of capillary tubes and found that the long-term longitudinal dispersion coefficient D_L was a power law of the seepage velocity with an exponent lower than 2 and approaching 1 in the most heterogeneous networks considered (see also Sahimi et al., 1986).

The theory of Taylor dispersion has been generalized in various ways. In particular, adsorption–desorption of the solute on the capillary walls was considered (e.g., Boddington and Clifford, 1983). Adsorption–desorption of the solute does not affect the flow of the solvent but may significantly retard the solute transport and enhance dispersion by allowing the solute particles to be trapped on the capillary walls. Although the published models (e.g., Boddington and Clifford, 1983) predict a Fickian long-term behavior, non-Fickian breakthrough curves were observed in laboratory experiments as well as in numerical simulations (Madras et al., 1996, and references therein) supporting the idea that hydrodynamic tracer traps (i.e., stagnant regions) may indeed cause non-Fickian dispersion (e.g., Berkowitz and Scher, 1995; Cortis and Berkowitz, 2004; Gouze et al., 2008; Levy and Berkowitz, 2003).

11.02.6 Multiphase Flow

In many geologic contexts, porous media are not saturated with a single fluid phase but instead contain mixtures of immiscible phases, for instance, water and air in the vadose zone or water, oil, and gas in hydrocarbon reservoirs. When two immiscible fluids are in contact, they form an interface that has an excess of free energy with respect to the bulk fluids. The two-phase interface is characterized by a surface energy density that depends on the specific phases in contact (e.g., 0.072, 0.048, 0.024, and 0.48 J m^{-2} for water–air, water–oil, air–oil, and air–mercury, respectively). Possessing free energy, a two-phase interface will spontaneously contract as in response to a surface tension. Surface tension is a useful concept for modeling the behavior of a two-phase interface (note that it has the same values and units as the surface energy density, although it is more appropriate to use N m^{-1} for a tension than J m^{-2}). The surfaces of solids are also subject to surface energy and surface tension, but, of course, surface tension is too weak, in this case, to produce any noticeable deformation. The forces associated with surface tension are called capillary forces. The dynamic behavior of the different fluids present in the Earth is thus governed not only by viscous, mechanical, and gravitational forces but also by capillary forces.

11.02.6.1 Wettability, Capillary Pressure, Drainage, and Imbibition

In static equilibrium, a liquid droplet immersed in gas and resting on a planar solid substrate looks like the pictures in Figure 6. The contact angle θ of the liquid–gas interface with the solid plane at a point along the line of contact reflects the balance of the surface tensions between solid and gas, γ_{gs}; liquid and solid, γ_{ls}; and gas and liquid, γ_{gl}, as expressed by Young's equation:

$$\gamma_{gs} - \gamma_{ls} = \gamma_{gl} \cos\theta \quad [40]$$

When θ is $<90°$ as in Figure 6(a), the liquid phase tends to expand over the solid surface and is said to be the wetting phase. In the other case (Figure 6(b)), it tends to contract and is called the nonwetting phase. However, the contact angle is a more complex property than idealized in Figure 6. It is affected by the motion of the fluids as well as by certain properties of the solid surface (e.g., roughness and presence of impurities or surfactants; Bico et al., 2002; Chow, 1998; de Gennes, 1985). Wettability is one of the most important properties controlling how immiscible fluids get distributed inside a porous material (e.g., Bear, 1972; Dullien, 1992; Sahimi, 1995).

Even in static equilibrium, pressure is not continuous across a curved interface separating two immiscible fluids (when the interface is flat, the pressures are equal). The capillary pressure (i.e., difference in pressure between two immiscible fluids) depends on the surface tension γ and the radius of curvature R_m of the interface according to Laplace's law, $P_{cap} = 2\gamma/R_m$. The fluid pressure is higher on the concave side of the curved interface, which, in the case of a porous medium, always corresponds to the nonwetting phase. It follows that pressure

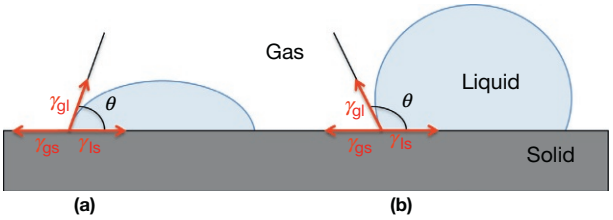


Figure 6 Two examples of the contact angle θ between two immiscible fluid phases (e.g., a liquid and a gas) and a solid. In case (a), θ is $<90°$ and the liquid is wetting. Case (b) shows a nonwetting liquid.

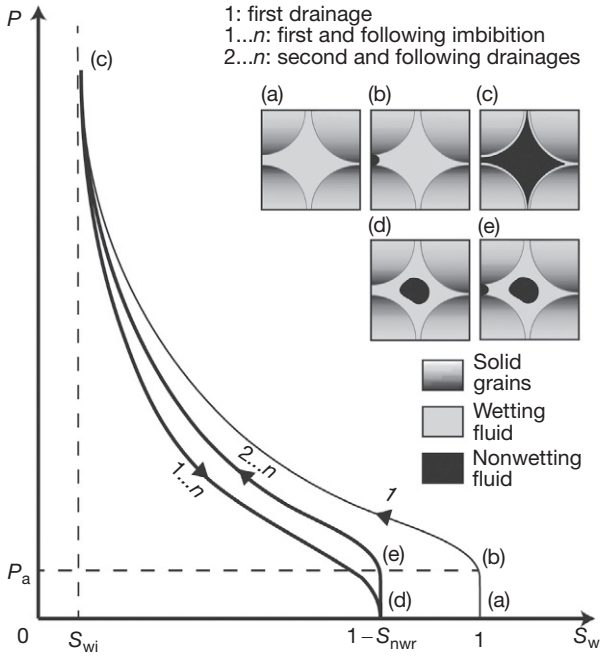


Figure 7 A generic example of capillary pressure curves (i.e., nonwetting fluid pressure plotted against wetting fluid saturation). During the first drainage, a nonwetting fluid (represented as black in the insets) is gradually injected into a porous body initially saturated with a wetting fluid (light gray). This stage is followed by imbibition–drainage cycles that all follow the same curves. The insets illustrate the pore-scale distribution of the phases at various stages during the drainage–imbibition cycles. (a) In the initial state, only the wetting fluid is present. (b) A small nonwetting fluid meniscus appears on the left side, indicating the onset of penetration by the nonwetting fluid. (c) The irreducible wetting fluid saturation is shown as a film of wetting fluid along the pore walls (implying that θ is zero). (d) At the residual nonwetting fluid saturation, some amount of nonwetting fluid is trapped in the pore space. (e) A small nonwetting fluid meniscus appears on the left side, indicating the onset of penetration by the nonwetting fluid at the beginning of a new drainage cycle.

must be applied to a nonwetting fluid in order to force it to penetrate a porous medium initially at total saturation of the wetting fluid and must be increased to produce further infiltration (a process called drainage, illustrated in Figure 7). If the pressure is released, the nonwetting fluid is spontaneously expelled, although a portion of it may remain trapped in the porous medium. During drainage, the saturation in the wetting fluid, $S_w = 1 - S_{nw}$, remains equal to 1 until the applied pressure reaches a critical value, the access pressure P_a, at which point the nonwetting fluid penetrates the largest pore throat(s) directly accessible on the surface of the porous body. As the pressure P of the nonwetting fluid is increased, smaller and smaller pore throats are penetrated and a greater volume of nonwetting fluid is injected, displacing a certain amount of wetting fluid depending on the compressibilities of the fluids. Finally, S_w asymptotically reaches the 'irreducible' saturation S_{wi}, corresponding to the amount of wetting fluid trapped in domains of the pore space nonaccessible to the nonwetting fluid (e.g., thin film along the surface of the grains as in Figure 7, in the microporosity, or even in macropores only accessible through extremely small throats). If the nonwetting fluid pressure is gradually lowered back to zero, the wetting fluid reinvades the porous medium. During this process (called imbibition), a different pressure–saturation curve is followed, situated under the initial drainage curve, and ending not at $S_w = 1$ but at $S_w = 1 - S_{nwr}$, where S_{nwr} denotes the so-called residual nonwetting fluid saturation (Figure 7). Thus, some amount of nonwetting fluid gets trapped in the porous medium at the end of the drainage–imbibition cycle. During the next drainage step, S_w decreases from $1 - S_{nwr}$ to S_{wi}, following a pressure–saturation curve that lies between the previous two (Figure 7). Subsequent imbibition–drainage cycles usually follow the hysteretic curve formed by the initial imbibition and the second drainage. In practice, the imbibition–drainage curves (often called capillary curves in the literature) are difficult to measure accurately because the imbibition–drainage processes can be strongly affected by dynamics and end effects (e.g., Bear, 1972). Although the imbibition–drainage processes are considered to occur in many geotechnical and natural situations (e.g., oil recovery by water flooding and infiltration of rain water in unsaturated soil), they are, by no means, the only processes causing changes in immiscible fluid saturations in porous rocks in the Earth. For example, a gas initially dissolved in the saturating pore liquid may come out of solution if the thermodynamic conditions are changed and bubbles of a second phase are formed (e.g., Zuo et al., 2011). The capillary pressure curves determined during imbibition–drainage cycles (Figure 7) have no equivalent in the case of exsolved gas bubbles since the gas phase is discontinuous.

11.02.6.2 Two-Phase Flow, Relative Permeabilities

The macroscopic equations describing two-phase flow in porous media can be obtained by the method of volume averaging (e.g., Whitaker, 1986b). Let us consider a porous medium REV containing two immiscible fluids, labeled 1 and 2. Each of the two phases verifies a continuity equation and an equation of motion:

$$\nabla \cdot \mathbf{u}_i = 0 \qquad [41]$$

and

$$\eta_i \nabla^2 \mathbf{u}_i = \nabla p_i \qquad [42]$$

where the subscript i refers to phase 1 or phase 2. In addition, the following internal boundary conditions must be satisfied: (i) no-slip boundary condition for either phase in contact with the solid, (ii) continuity of the phase velocities across the two-phase interface, and (iii) continuity of the stresses in each phase across the interface:

$$-p_1 \mathbf{n}_{12} + \mathbf{t}_1 \cdot \mathbf{n}_{12} = -p_2 \mathbf{n}_{12} + \mathbf{t}_2 \cdot \mathbf{n}_{12} + 2\gamma_{12} R_{12} \mathbf{n}_{12} \qquad [43]$$

where $\mathbf{n}_{12}$ is a unit vector normal to the interface, R_{12} is the mean curvature of the interface, γ_{12} is the surface tension, and the total stress tensor in each phase was decomposed as $-p_i \mathbf{I} + \mathbf{t}_i$. By a method similar to the one used for establishing Darcy's law (see Section 11.02.4), that is, applying volume averaging to the equations and boundary conditions mentioned earlier, solving the relevant closure problem, and using a number of simplifying assumptions (in particular, that the mean curvature of the two-phase interface is not

affected by the motion of the fluids), Whitaker (1986b) arrived at the following macroscopic equations:

$$\langle \mathbf{u}_i \rangle = -\frac{\mathbf{K}_i}{\eta_i} \cdot \nabla \langle p_i \rangle_i + \mathbf{K}_{ij} \langle \mathbf{u}_j \rangle \qquad [44]$$

where the tensor $\mathbf{K}_i$ can be understood as a permeability tensor specific to phase i and $\mathbf{K}_{ij}$ accounts for the viscous drag of one phase upon the other. The second term on the right-hand side of eqn [44] is considered to be negligible in many situations of interest (if there is a large viscosity contrast between the phases (Whitaker, 1986b) or if the contact angle is such that the phases tend to settle in different pores and therefore do not interact with each other (de Gennes, 1983)). However, this point still needs some clarification today because lattice Boltzmann numerical simulations produce significant viscous drag terms $\mathbf{K}_{ij}$, at least when the de Gennes condition is not satisfied (e.g., Rothman, 1990). Neglecting the viscous drag terms and assuming strict conditions of isotropy of both the medium and the distribution of the phases, we obtain the classic two-phase flow equations:

$$\mathbf{q}_i = -\frac{k_i}{\eta_i} \nabla P_i = -\frac{k_i^r}{k_0 \eta_i} \nabla P_i \qquad [45]$$

where the subscript i stands for w or nw, the wetting and nonwetting phases, respectively; k_0 is the absolute (i.e., single-phase flow) permeability; and k_i^r denotes the so-called relative permeabilities. Equation [45] has often been criticized in the past, in particular regarding the implication that the relative permeabilities k_w^r and k_{nw}^r uniquely depend on the saturations of the phases (e.g., Bear, 1972). The relative permeabilities, k_w^r and k_{nw}^r, are usually measured during an imbibition–drainage cycle and are plotted as a function of the saturation S_w (see Figure 8, where the hysteresis associated with the imbibition–drainage cycle was not represented to simplify the diagram). It is usually observed that k_w^r and k_{nw}^r are monotonic functions of S_w, declining sharply from their maxima and vanishing at wetting fluid saturations of S_{wi} and $1-S_{nwr}$, respectively. The sum $k_w^r + k_{nw}^r$ is significantly lower than 1 at intermediate saturations and depends on the viscosity ratio η_w/η_{nw}, implying that strong phase interference occurs when both phases are mobile (e.g., Persoff and Pruess, 1995). A large amount of work has been expended to model the relative permeability curves (Figure 8) and to link them to capillary curves (Figure 7). A multitude of models has been published and it would be difficult for us to give a full account of all of them. They generally attempt to model relative permeabilities as sole functions of saturations. Even though this working assumption appears to be adequate for all practical purposes (e.g., Bear, 1972), it cannot be true in general. Recent experimental work shows that relative permeabilities depend on other factors such as the viscosity ratio, the capillary number (a dimensionless number measuring the relative magnitudes of the viscous and capillary forces), and even the flow rates (e.g., Avraam and Payatakes, 1995; Tallakstad et al., 2009). The shortcomings of the classic models are apparent when they are simultaneously applied to a single data set and their results compared (e.g., Chen et al., 1999; Li and Horne, 2006). The fitting quality of the models tends to be somewhat erratic and, furthermore, it often happens that a model satisfactorily fitting the data for one of the phases produces large discrepancies for the other one (e.g., Li and Horne, 2006). In such a case, one can reasonably conclude that the good fit was mere coincidence and not the result of sound physics. The classic models also may not be appropriate for situations in which a process other than imbibition–drainage occurred (e.g., Zuo et al., 2011, found that, during gas exsolution, the relative permeabilities were much smaller than the imbibition–drainage ones at equivalent saturations). One concept that seems to gain acceptance recently is that of flow morphology. Flow morphology changes drastically depending on the saturation and when the viscosity ratio, capillary number, and flow rates are changed (e.g., Avraam and Payatakes, 1995). Among the flow morphologies identified, several are characterized by a discontinuous and nevertheless mobile nonwetting phase, with formation and motion of nonwetting ganglia of various sizes. In such cases, it is clear that the viscous interactions between the phases must be large. Only at high capillary numbers and high nonwetting saturations do the phases form continuous, parallel paths, as is implicitly assumed by eqn [44] and the classic relative permeability models.

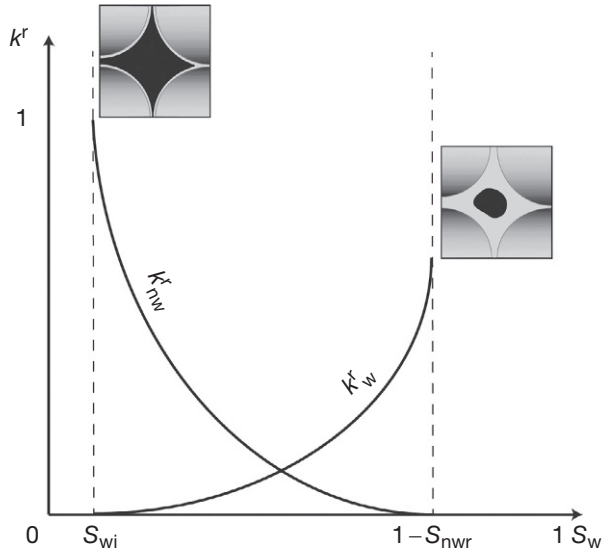


Figure 8 A generic example of relative permeability curves. The insets illustrate the distribution of the phases in the pore space at the end points.

11.02.7 Processes Affecting Permeability

Because permeability is so sensitive to the geometric details of the pore space, any process affecting the pore geometry is bound to alter permeability and fluid flow. Natural processes occurring in sediments and porous rocks are usually complex composite processes, resulting from the combination of simultaneous, possibly coupled, elementary mechanisms (e.g., see the discussions of the diagenesis of sedimentary rocks, cited in Section 11.02.2). Several mechanisms of porosity change in granular materials were discussed by Evans et al. (1999) and are summarized in Figure 9. These mechanisms operate in response to a variety of mechanical and/or chemical driving forces. Among them, microfracturing, frictional sliding,

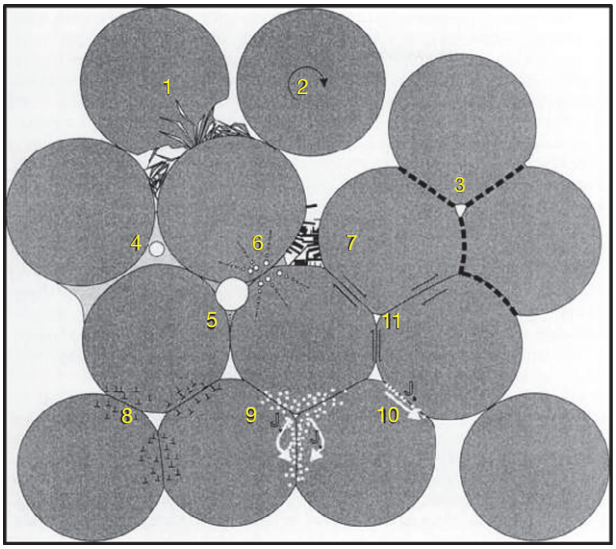


Figure 9 An idealized illustration of several mechanisms of porosity change possibly active in granular aggregates, depending on the local conditions of temperature, fluid pressure, applied stresses, and chemical potentials of various species (courtesy Evans et al., 1999): (1) Hertzian microcracking, (2) grain rearrangement by frictional sliding, (3) pressure solution, (4) precipitation of a new mineral phase, (5) overgrowth, (6) crack healing, (7) diagenetic or metamorphic reactions, (8) crystal plasticity via dislocation motion, (9) lattice diffusion creep, (10) grain boundary diffusion creep, and (11) grain boundary sliding (superplasticity).

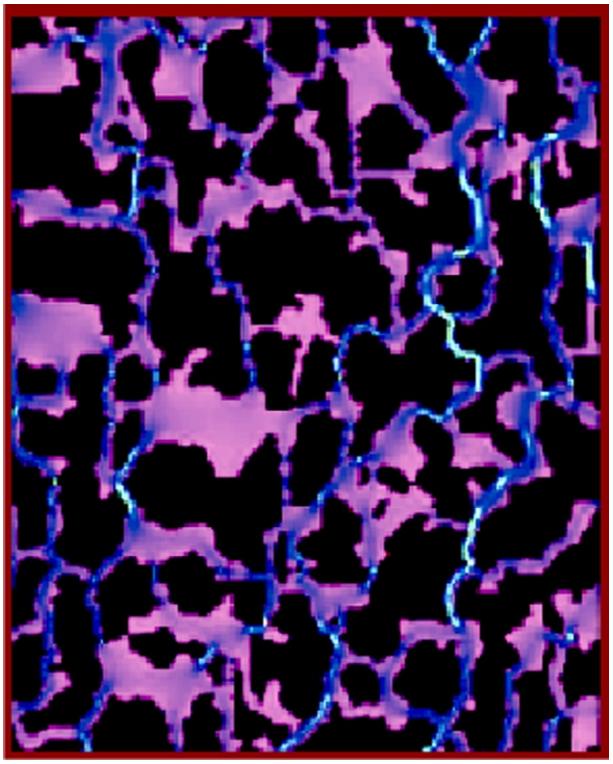


Figure 10 An example of fluid velocity field obtained by numerical simulation of flow through a two-dimensional porous medium based on microstructure images of Berea sandstone (courtesy Boek and Venturoli, 2010). In this picture, the solid phase is represented in black and the magnitude of the local fluid velocity is shown according to a continuous color scale, representing the points of near zero velocities as reddish-purple pixels and the maximum velocities as bluish-white ones. The predominantly blue and purple regions correspond to the effective and noneffective porosity, respectively (see text for more explanations). In this Berea sandstone example, the effective porosity is predominantly associated with the presumably easily compressed pore constrictions (or throats).

pressure solution, dissolution/precipitation, and diagenetic reactions occur at temperatures lower than about a third of the rock melting point T_m, while metamorphic reactions require somewhat higher temperatures, and dislocation creep and diffusion creep only produce significant strain rates at temperatures well in excess of $T_m/3$. Pressure solution, dissolution–precipitation, and diagenetic reactions need the presence of a solvent fluid, and their results may depend strongly on the fluid flow rate. Needless to say, this list is far from complete, especially if materials more complex than granular media are considered.

Processes involving some of these mechanisms were investigated in laboratory experiments (for a review, see Bernabé et al., 2003, and references therein). Simultaneously recording porosity and permeability during such experiments allows construction of porosity–permeability relationships (PPRs). A 'universal,' one-to-one PPR valid for all porous media cannot exist (e.g., Bernabé et al., 2003), but the experimental PPRs mentioned earlier are material- and process-specific and therefore uniquely defined. One convenient way to characterize the PPRs is to cross plot porosity and permeability data in log–log scale. The porosity–permeability curves are thus represented by

$$\frac{k}{k_0} = \left(\frac{\phi}{\phi_0}\right)^{\xi(\phi)} \qquad [46]$$

where $\xi(\phi)$ is simply equal to the local slope of the $\log(k)$–$\log(\phi)$ curve. In this approach, the exponent $\xi(\phi)$ is a measure of the (local) sensitivity of permeability to changes in porosity. Bernabé et al. (2003) proposed to relate $\xi(\phi)$ to changes in the hydraulic effectiveness of the pore space. As discussed in **Section 11.02.5**, convection transport is dominant in some portions of the pore space during fluid flow, while the rest is nearly stagnant (e.g., see **Figure 10** for an illustration). Effective porosity can then be defined as the volume fraction of the former and noneffective porosity of the latter. A low value of the exponent ξ (i.e., $\xi \ll 1$) indicates that permeability was essentially unaffected by the porosity change and therefore suggests that noneffective porosity was predominantly destroyed (or created in the case of a porosity-creating mechanism). Obviously, high values (i.e., $\xi \gg 1$) imply the reverse. Bernabé et al. (2003) showed that neutral processes, which do not change the relative amounts of effective and noneffective porosity, should be associated to ξ's in the range of 1–3 (a more precise value cannot be assigned).

11.02.7.1 Examples of Process-Specific PPRs

11.02.7.1.1 Elastic compression

Among the mechanical processes affecting rocks (see Bernabé et al., 2003), the most basic one is elastic compression.

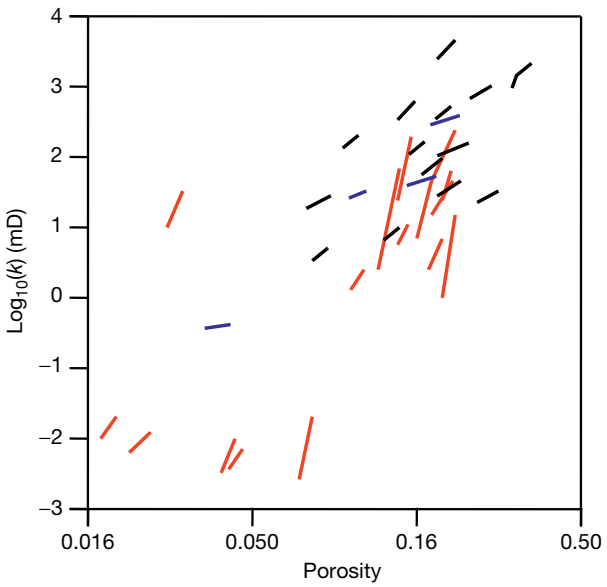


Figure 11 Schematic representation of the porosity–permeability relationships (PPRs) compiled by Bernabé et al. (2003) for various sandstones undergoing elastic compression. Note that the original data points are not represented here, but only the trends derived from them. As a help for visualization, different ranges of the exponent ξ_e are represented with different colors, namely, red, black, and blue for $\xi_e > 5$, $5 > \xi_e > 2$, and $2 > \xi_e > 1$, respectively. The exponent ξ_e does not correlate with either ϕ or k.

Surprisingly, this presumably simplest of all processes produces an extremely wide range of ξ's in different sandstones (**Figure 11**). David et al. (1994) measured the exponent ξ in the elastic regime (thereafter denoted ξ_e) in several sandstones and reported values as high as 25, whereas other sandstones compiled in Bernabé et al. (2003) had ξ_e's approaching one. Based on Bernabé et al.'s (2003) interpretation of ξ, it can be concluded that, in the rocks with the highest ξ_e's, the most compressible pores (or portions of pores) must also be hydraulically effective, whereas it is the stagnant pores that should be most deformable in rocks with the lowest ξ_e's. Clearly, these two end-members correspond to different microstructures. For instance, the pore-throat microstructure illustrated in **Figure 10** may occur in some high ξ_e sandstones. But, it is not as easy to imagine what kind of microstructure gives rise to low ξ_e's. Of course, several different microstructures may correspond to the same value of ξ_e, especially if rocks other than sandstones are considered. The wide range of ξ_e's shown in **Figure 11** is particularly remarkable because sandstones do not have as much microstructural variability as, for example, carbonates or volcanic rocks.

Understanding the origin of a particular value of ξ_e, or, in other words, identifying the cause of the pressure sensitivity of permeability, may require an extremely detailed characterization of the microstructure. For example, it is tempting to apply the fracture permeability models of Gangi (1978) and Walsh (1981) to rocks predominantly containing open microcracks or crack-like grain boundaries. However, these models rely on a statistical characterization of the fracture roughness that cannot be accurately carried out at the scale of the microcrack aperture (< 1 μm). In these models, the pressure dependence of permeability is nonlinear and decelerating since rough cracks stiffen under increasing pressure. An alternative model assuming a heterogeneous population of smooth microcracks with variable aspect ratios (i.e., ratio of the short over the long crack dimensions) was proposed by Gavrilenko and Guéguen (1989). Under increasing pressure, an increasing fraction of the crack population closes (i.e., the microcracks with the lowest aspect ratios), causing the crack network to lose connectivity gradually. At some finite pressure value, the percolation threshold is reached and permeability vanishes completely, an outcome that is not supported by experimental evidence (the permeability of cracked rocks generally remains finite even at very high pressures). The basic model was then modified to include a population of rough microcracks that is capable of preserving a finite permeability at very high pressures. Again, Gavrilenko and Guéguen's (1989) model requires extensive and detailed statistical information on the crack geometry that is still unavailable today.

11.02.7.1.2 Nonelastic compression

When sufficiently high temperatures and/or pressures are applied, the rock solid matrix experiences pervasive plastic deformation, leading to irreversible porosity loss. Examples of PPR observed during hot isostatic pressing of monomineralic granular aggregates (calcite (Bernabé et al., 1982) and quartz (Wark and Watson, 1998)) are plotted in **Figure 12** (note that arrows are now needed to visualize the irreversible rock evolution). In the case of calcite, the PPR first shows a moderate decrease of k at porosities above about 10% (i.e., exponent $\xi \cong 3$), followed by an accelerating permeability drop (i.e., steadily increasing ξ), and eventually a total vanishing of k at ϕ near 4%. In these experiments, Bernabé et al. (1982) also measured the unconnected porosity ϕ_u (a quantity that remained near zero at porosities larger than 10% and gradually increased to become equal to the total porosity at the percolation threshold ϕ_{thr} near 4%). Plotting k against the connected porosity, $\phi_c = \phi - \phi_u$, in log–log scale, revealed an approximate power law relationship with an exponent of 3:

$$k \propto (\phi - \phi_u)^3 \qquad [47]$$

which seamlessly continues the high porosity, $k \propto \phi^3$, segment described earlier. The accelerating permeability drop can therefore be understood as primarily caused by the gradual disconnection of the pore space occurring at porosities below 10%. Mavko and Nur (1997) modified the Kozeny–Carman model to account for the occurrence of a percolation threshold and applied it to Bernabé et al.'s (1982) data. They found that the cubic function

$$k \propto (\phi - \phi_{thr})^3 \qquad [48]$$

(with $\phi_{thr} = 4.5\%$) also produced an excellent fit to the experimental data (but it must be noted that the expressions [47] and [48] are not mathematically equivalent since ϕ_u is not a constant). In the case of quartz, pore disconnection did not occur, even at very low porosities, and simple power law PPRs with ξ between 2.5 and 3 were measured. This result is not consistent with Bourbié and Zinszner's (1985) ϕ–k data on the nearly 100% pure quartz Fontainebleau sandstone, which

indicate the existence of a percolation threshold for this rock (the data were well fitted by eqn [48] with $\phi_{thr}=2.5\%$). More complex, calcite–quartz aggregates were also studied (Xiao et al., 2006). Since the dominant mineral phase was calcite, the results were similar to those obtained for pure calcite aggregates. Interestingly, it was observed that the percolation threshold ϕ_{thr} increased with increasing quartz content and that permeability decreased at much slower rates during creep under deviatoric stresses than in isostatic pressure conditions.

At room temperature, halite aggregates submitted to isostatic compression display semibrittle behavior (i.e., partially plastic and partially brittle; see Bernabé et al., 2003, and references therein). The corresponding PPRs are nearly straight lines in log–log scale with ξ between 5 and 7 (Spangenberg et al., 1998). The cataclastic (distributed brittle deformation) compression experiments on various sandstones of Zhu and Wong (1997) display even larger ξ's. These observations suggest that brittle mechanisms are more prone to destroy effective porosity than plastic ones. The results of room temperature, ductile consolidation experiments on mudstones (Dewhurst et al., 1998) and well-casing cement (Ghabezloo et al., 2009) were also included in **Figure 12**. These are unusual ultrafine-grained materials characterized by very low permeabilities and very high porosities. The associated PPRs displayed surprisingly high ξ's. The origin of these large values of ξ is difficult to assess since the grain-scale mechanisms responsible for the deformation are almost impossible to observe but the PPRs do not seem to indicate the approach to a percolation threshold.

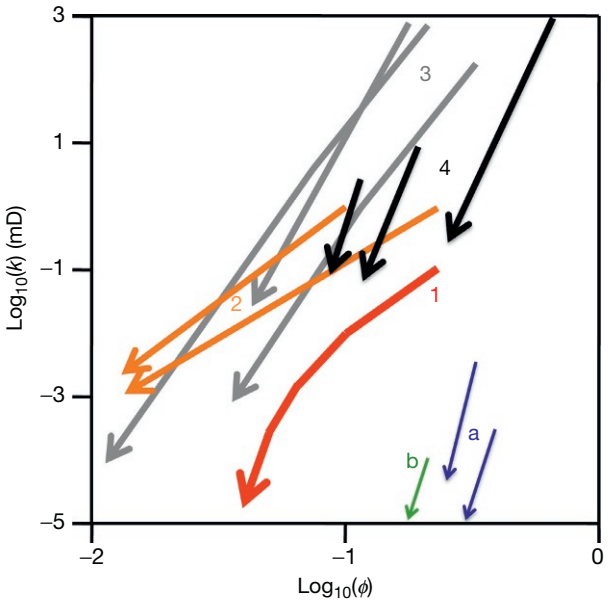


Figure 12 Schematic representation of the PPRs for granular materials undergoing different types of nonelastic compression: hot isostatic pressing of monomineralic aggregates, (1) calcite (Bernabé et al., 1982) and (2) quartz (Wark ans Watson, 1998); (3) semibrittle compression of halite aggregates (Spangenberg et al., 1998); (4) cataclastic compression of sandstones (Zhu and Wong, 1997); and ductile consolidation of (a) mudstones (Dewhurst et al., 1998) and (b) well-casing cement (Ghabezloo et al., 2009). For more details on 1–4, see Bernabé et al. (2003). As in **Figure 11**, the original data points are not represented here, but only the trends derived from them.

Experiments were also conducted with added deviatoric stresses, but the resulting permeability anisotropy and the possible occurrence of dilatance and strain localization complicate the interpretation (e.g., Zhu and Wong, 1997). Zhu et al. (2002, 2007a) proposed a damage model (i.e., microcracking and pore collapse) of stress-induced permeability anisotropy during distributed cataclastic flow. Dilatant strain localization has also been extensively studied using recording and analysis of acoustic emissions (e.g., Lockner et al., 1991; Stanchits et al., 2006).

Gradual, distributed microcracking was separately explored in a different set of experiments (e.g., Zhang et al., 1994, 2001, and other references in Bernabé et al., 2003). In these studies, microcracks were mechanically and/or thermally induced in initially fully dense, impermeable rocks, thus producing a simultaneous increase in porosity and permeability. Typical dilatant microcracking PPRs start at porosities of a few tens of percent with a very steep increase in permeability (i.e., very high ξ) and bend toward much lower slopes when the porosity exceeds a few percent. This shape is typical of a percolation process controlled by changes in crack connectivity and can be well accounted for by the permeability model of Guéguen and Dienes (1989), which includes a factor varying from 0 to 1, expressing the probability for a crack to belong to the percolating cluster. When microcracks are introduced into an initially porous and permeable material, this percolation factor remains nearly constant, close to 1, explaining why permeability is only moderately affected in this case.

11.02.7.1.3 Dissolution–precipitation, alteration, and chemical reactions

The greatest changes in porosity occurring in Earth's crust are produced by chemical fluid–rock interactions (see the references on diagenesis in **Section 11.02.2**). By circulating super- or undersaturated fluids through a rock specimen, the effect of precipitation or dissolution of a particular solid can be studied in the laboratory (e.g., Colón et al., 2004; Luquot and Gouze, 2009; McCune et al., 1979; Reis and Acock, 1994; Zhang et al., 2001). Examples of precipitation and dissolution PPRs are shown in **Figure 13** as downward- and upward-oriented arrows, respectively. They reveal a wide variety of behavior depending on the rock considered (sandstone or carbonate) and the experimental conditions used (flow rate, fluid composition, pH, and so forth). In some cases, very high ξ values were observed, while, in others, ξ significantly lower than 1 was found. A few experiments showed a spectacular change in behavior (from low to high ξ or the reverse), possibly caused, in the case of the dissolution experiments, by the formation of preferential conduits or 'wormholes' (Luquot and Gouze, 2009). To some extent, the experimental conditions can be characterized by the dimensionless Péclet and Damköhler numbers (Pe and Da, respectively), which allow for a comparison of the reaction and transport time rates. Luquot and Gouze (2009) estimated ξ and Da during their experiments but could not find a simple interrelationship, perhaps because they were not able to explore a sufficiently large range of Damköhler number (they used initial values of Da between 1 and 6, but the Da's rapidly decreased in the course of the experiments and ended up in a narrow range, from infinitesimal values up to 2).

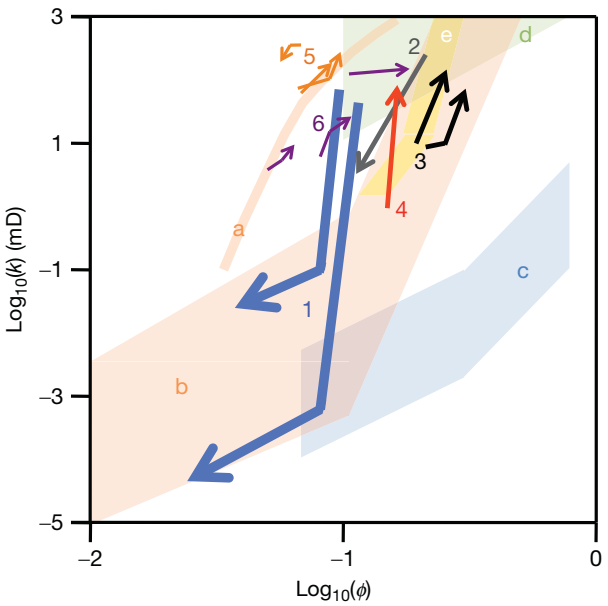


Figure 13 Schematic representation of the PPRs associated with flow of variously reactive fluids through porous materials and rocks: (1) chemical alteration of sintered silica glass beads (Mok et al., 2002); (2) precipitation in Berea sandstone (Reis and Acock, 1994); acidization of (3) phacoides sandstone and (4) spergen limestone (McCune et al., 1979); (5) precipitation and dissolution in Mondeville oolitic limestone (Luquot and Gouze, 2009); and (6) dissolution in Fontainebleau sandstone (Colón et al., 2004). As in **Figure 11**, the original data points are not represented here, but only the trends derived from them. In addition, several data sets for natural rocks are sketched in this diagram as lightly shaded ϕ–k regions: (a) Fontainebleau sandstone (Bourbié and Zinszner, 1985); (b) shales and sandstones (Pape et al., 2000; Revil and Cathles, 1999); (c) fine-grained chalks (Mallon et al., 2005) and subduction zone sediments (Gamage et al., 2011); (d) vesicular basalts (Saar and Manga, 1999) and degassed lavas (Mueller et al., 2005); and (e) seafloor vent deposits (Zhu et al., 2007b).

Mok et al. (2002) induced chemical alteration of silica glass beads sintered to porosities between 8% and 15%. They observed unusual PPRs, starting with a rapid decline of permeability by several orders of magnitude (i.e., $\xi > 10$) and, after a sharp transition, ending with a much weaker decrease of k with ξ near 2 (see simplified sketches of these PPRs in **Figure 13**). According to Mok et al. (2002), the alteration reactions during the first stage transformed the initially very smooth glass–fluid interfaces into increasingly rough pore walls, hence causing a much stronger reduction in permeability than in porosity. Once a significant roughness level was reached, permeability became insensitive to additional roughening and decreasing pore dimensions were then primarily responsible for the subsequent permeability reduction (corresponding to low values of ξ).

Another important class of reaction-controlled phenomena concerns the dehydration of minerals such as serpentinite. The release of water associated with dehydration reactions produces a transitory rise in pore pressure and a general weakening of the rock (e.g., Ko et al., 1997; Tenthorey and Cox, 2003). These changes are accompanied by increases of porosity, pore connectivity, and permeability.

11.02.7.2 Natural Processes, Diagenesis

Permeability and porosity data have been collected for rocks and other geologic materials that acquired their present properties via natural processes similar to the laboratory processes discussed previously. These data are grouped according to the rock families investigated. For example, a partial list of rock types may include mudstones (Dewhurst et al., 1998; Yang and Aplin, 2010), clays and shales (Pape et al., 2000; Revil and Cathles, 1999), sands and sandstones (Bosl et al., 1998; Bourbié and Zinszner, 1985; Cade et al., 1994; Kameda et al., 2006; Nelson, 2000; Pape et al., 2000; Revil and Cathles, 1999), carbonates (Ehrenberg et al., 2006), chalks (Mallon et al., 2005), sediments in subduction zones (Gamage et al., 2011), effusive volcanic rocks (Mueller et al., 2005; Saar and Manga, 1999), and seafloor vent deposits (Gribbin et al., 2012; Zhu et al., 2007b). Undoubtedly, different materials from this list did not necessarily experience the same history even if they belonged to the same rock family. They were most likely not subjected to the same processes and, therefore, should not verify the same PPRs. Even the ϕ–k trend, beautifully delineated in the Fontainebleau sandstone data set of Bourbié and Zinszner (1985), may not represent a true process-specific PPR, comparable to the experimental ones discussed in the previous subsections.

For the sake of comparison with the experimental data described in **Section 11.02.7.1**, regions of the ϕ–k space covered for some of the natural rocks cited earlier were highlighted in **Figure 13**. The shale and sandstone data of Revil and Cathles (1999) and Pape et al. (2000) lie in a wide zone with a similar hockey-stick shape as the PPRs for altered glass beads measured by Mok et al. (2002). Remarkably, the Fontainebleau sandstone data set is positioned entirely outside this zone and has a quite different shape, reminiscent of the hot-pressed calcite PPR discussed earlier, including the likely presence of a percolation threshold. However, the uncommon behavior of Fontainebleau sandstone will stand out less clearly if more sandstone data (e.g., Nelson, 2000) were added to **Figure 13**. The Nelson (2000) data sets include examples of both the 'percolation-type' PPR characteristic of Fontainebleau sandstone and the (hockey-stick-shaped) 'roughness-type' PPR typical of shales and clay-rich sandstones. The former kind is more frequently associated with quartz arenites, while the latter corresponds to feldspathic litharenites (note that feldspar is, indeed, more easily altered in hydrothermal conditions than quartz). Using a numerical simulation approach, Cade et al. (1994) submitted a dense pack of identical spheres to various processes, such as 'compaction' (i.e., the reduction of the intergranular distances with interpenetration of the spheres allowed) or 'cementation' (i.e., the increase of the sphere radius with interpenetration again allowed). All the processes investigated yielded percolation-type PPRs with a percolation threshold value depending on the process considered (see Cade et al., 1994, for more details; similar results were obtained by Keehm et al., 2001). Hence, it appears plausible that percolation-type PPRs are caused by a combination of mechanical compaction and cementation, whereas roughness-type PPRs may be produced by dissolution/alteration of feldspars and/or growth of aluminosilicates in the pore space.

Zhu et al. (2007b) also contended that a roughness-type PPR evolution was likely in seafloor vent deposits but attributed it to connectivity loss in the early stage of chimney growth (Figure 13). Side flanges of seafloor vents were found to have a layered structure that produced highly anisotropic PPRs with ξ on the order of 1 or 2 for k parallel to the layers and ξ values near 9 in the perpendicular direction (Gribbin et al., 2012).

Figure 13 also illustrates the absence of a general PPR for natural materials. For example, consider the vesicular basalts and degassed lavas (Mueller et al., 2005; Saar and Manga, 1999) and compare them to the ultrafine, low-permeability sediments of Mallon et al. (2005) and Gamage et al. (2011). These materials cover essentially the same range of porosities but nevertheless differ by 4 orders of magnitude in permeability.

11.02.7.3 Summary

The permeability and porosity of rocks in the Earth crust are dynamic quantities varying with time and space. They are affected by a variety of coupled mechanical and chemical processes. Some of these processes have been studied in the laboratory and the corresponding PPRs measured. One convenient way to characterize the PPRs is to use power laws (i.e., $k \propto \phi^{\xi}$) with a variable exponent ξ that can be interpreted as a measure of the local change in the proportions of hydraulically effective and noneffective pore space. In the Earth, multiple porosity-changing mechanisms may be simultaneously activated depending on the rock microstructure and the in situ physical conditions. The rock evolution may, in turn, induce changes of the physical conditions and the activated mechanisms. Such feedback loops are very difficult to model and their results very difficult to predict.

11.02.8 Concluding Statement

In this chapter, we tried to argue that the main goal of rock physics is to upscale the phenomena or processes of interest from the pore scale to the relevant macroscopic scale (i.e., the core scale in the laboratory or the reservoir scale in field studies). In order to obtain an accurate macroscopic description, a detailed knowledge and a precise understanding of the phenomena at the microscopic scale are necessary. Thanks to new technological advances in 3D microimaging and numerical algorithms, very complex, coupled chemomechanical pore-scale processes can now be simulated on the computer, allowing direct testing of the upscaling methods and results. However, when attempting to simulate natural phenomena, the appropriateness and/or reliability of particular numerical methods may be hard to assess. Laboratory experiments therefore remain an essential element in our effort to establish a firm pore-scale basis to our macroscopic models.

Acknowledgments

We are grateful to Lee Slater, Jörg Renner, and an anonymous reviewer for many helpful suggestions. YB's work is partially supported by the US Department of Energy under grant DE-FG09-97ER14760. This is IPGP contribution no. 3426.

References

Adler PM (1992) *Porous Media: Geometry and Transport,* 544 pp. Boston, MA: Butterworth-Heinemann.

Adler PM, Jacquin CG, and Quiblier JA (1990) Flow in simulated porous media. *International Journal of Multiphase Flow* 16: 691–712.

Adler PM and Thovert JF (1999) *Fractures and Fracture Networks*. Dordrecht: Kluwer, 433 pp.

Aharonov E and Rothman DH (1993) Non-Newtonian flow (through porous media): A lattice-Boltzmann method. *Geophysical Research Letters* 20: 679–682.

Al-Kharusi AS and Blunt MJ (2007) Network extraction from sandstone and carbonate pore space images. *Journal of Petroleum Science and Engineering* 56: 219–231.

Al-Raoush RI and Willson CS (2005) Extraction of physically realistic pore network properties from three-dimensional synchrotron X-ray microtomography images of unconsolidated porous media systems. *Journal of Hydrology* 300: 44–64.

Arns CH, Knackstedt MA, Pinczewski WV, and Garboczi EJ (2002) Computation of linear elastic properties from microtomographic images: Methodology and agreement between theory and experiment. *Geophysics* 67: 1396–1405.

Arns CH, Knackstedt MA, Pinczewski WV, and Lindquist WB (2001) Accurate estimation of transport properties from microtomographic images. *Geophysical Research Letters* 28: 3361–3364.

Arns CH, Knackstedt MA, Pinczewski WV, and Martys NS (2004a) Virtual permeametry on microtomographic images. *Journal of Petroleum Science and Engineering* 45: 41–46.

Arns CH, Robins V, Sheppard AP, Sok RM, Pinczewski WV, and Knackstedt MA (2004b) Effect of network topology on relative permeability. *Transport in Porous Media* 55: 21–46.

Aste T and Weaire D (2008) *The Pursuit of Perfect Packing*, 2nd edn., 200 pp. Boca Raton, FL: Taylor and Francis.

Auriault JL (2002) Upscaling heterogeneous media by asymptotic expansions. *Journal of Engineering Mechanics* 128: 817–822.

Auriault JL, Borne L, and Chambon R (1985) Dynamics of porous saturated media, checking of the generalized law of Darcy. *Journal of the Acoustical Society of America* 77: 1641–1650.

Auzerais FM, Dunsmuir J, Ferreol BB, et al. (1996) Transport in sandstone: A study based on three dimensional microtomography. *Geophysical Research Letters* 23: 705–708.

Avraam DG and Payatakes AC (1995) Flow regimes and relative permeabilities during steady-state two-phase flow in porous media. *Journal of Fluid Mechanics* 293: 207–236.

Baldwin CA, Sederman AJ, Mantle MD, Alexander P, and Gladden LF (1996) Determination and characterization of the structure of a pore space from 3D volume images. *Journal of Colloid and Interface Science* 181: 79–92.

Bear J (1972) *Dynamics of Fluids in Porous Media*, 764 pp. New York: Dover.

Bentz DP and Martys NS (1994) Hydraulic radius and transport in reconstructed model three-dimensional porous media. *Transport in Porous Media* 17: 221–238.

Berkowitz B (2002) Characterizing flow and transport in fractured geological media: A review. *Advances in Water Resources* 25: 861–884.

Berkowitz B and Scher H (1995) On characterization of anomalous dispersion in porous and fractured media. *Water Resources Research* 31: 1461–1466.

Bernabé Y (1998) Streaming potential in heterogeneous networks. *Journal of Geophysical Research* 103: 20827–20841.

Bernabé Y (2009) Oscillating flow of a compressible fluid through deformable pipes and pipe networks: Wave propagation phenomena. *Pure and Applied Geophysics* 166: 1–26.

Bernabé Y, Brace WF, and Evans B (1982) Permeability, porosity, and pore geometry of hot-pressed calcite. *Mechanics of Materials* 1: 173–183.

Bernabé Y, Li M, and Maineult A (2010) Permeability and pore connectivity: A new model based on network simulations. *Journal of Geophysical Research* 115: B10203. http://dx.doi.org/10.1029/2010JB007444.

Bernabé Y, Mok U, and Evans B (2003) Permeability–porosity relationships in rocks subjected to various evolution processes. *Pure and Applied Geophysics* 160: 937–960.

Bernabé Y, Zamora M, Li M, Maineult A, and Tang YB (2011) Pore connectivity, permeability and electrical formation factor: A new model and comparison to experimental data. *Journal of Geophysical Research* 116: B11204. http://dx.doi.org/10.1029/2011JB008543.

Bernal JD and Finney JL (1967) Random close-packed hard-sphere model, II: Geometry of random packing of hard spheres. *Discussions of the Faraday Society* 143: 62–69.

Berryman JG (1995) Mixture theories for rock properties. In: Ahrens TJ (ed.) *Rock Physics & Phase Relations: A Handbook of Physical Constants*, *American Geophysical Union*, pp. 205–228. Washington DC: American Geophysical Union.

Berryman JG (2005) Comparison of upscaling methods in poroelasticity and its generalizations. *Journal of Engineering Mechanics* 31: 928–936.

Bico J, Thiele U, and Quéré D (2002) Wetting of textured surfaces. *Colloids and Surfaces, A: Physicochemical and Engineering Aspects* 206: 41–46.

Biot MA (1962) Mechanics of deformation and acoustic propagation in porous media. *Journal of Applied Physics* 33: 1482–1498.

Biswal B, Manwart C, Hilfert R, Bakke S, and Øren PE (1999) Quantitative analysis of experimental and synthetic microstructures for sedimentary rocks. *Physica A* 273: 452–475.

Biswal B, Øren PE, Held RJ, Bakke S, and Hilfert R (2009) Modeling of multiscale porous media. *Image Analysis and Stereology* 28: 23–34.

Bjørlykke K, Ramm M, and Saigal GC (1989) Sandstone diagenesis and porosity modification during basin evolution. *Geologische Rundschau* 78: 243–268.

Blair SC, Berge PA, and Berryman JG (1996) Using two-point correlation functions to characterize microgeometry and estimate permeabilities of sandstones and porous glass. *Journal of Geophysical Research* 101: 20359–20375.

Bleam WF (1993) Atomic theories of phyllosilicates: Quantum chemistry, statistical mechanics, electrostatic theory and crystal chemistry. *Reviews of Geophysics* 31: 51–73.

Blunt MJ, Jackson MD, Piri M, and Valvatne PH (2002) Detailed physics, predictive capabilities and macroscopic consequences for pore-network models of multiphase flow. *Advances in Water Resources* 25: 1069–1089.

Boddington T and Clifford AA (1983) The dispersion of a reactive species (atomic hydrogen) in a flowing gas. *Proceedings of the Royal Society of London A* 389: 179–196.

Boek ES and Venturoli M (2010) Lattice-Boltzmann studies of fluid flow porous media with realistic rock geometries. *Computers & Mathematics with Applications* 59: 2305–2314.

Boitnott GN (1997) Use of complex pore pressure transients to measure permeability of rocks. In: Omega V (ed.) *Formation Evaluation and Reservoir Geology. Society of Petroleum Engineers*, 37–45, paper 38717, Society of Petroleum Engineers AIME.

Bosl WJ, Dvorkin J, and Nur A (1998) A study of porosity and permeability using a lattice-Boltzmann simulation. *Geophysical Research Letters* 25: 1475–1478.

Bourbié T and Zinszner B (1985) Hydraulic and acoustic properties as a function of porosity in Fontainebleau sandstone. *Journal of Geophysical Research* 90: 11524–11532.

Bourgeat A (1997) Two-phase flow. In: Hornung U (ed.) *Homogenization and Porous Media*, pp. 95–128. Berlin: Springer.

Boutt DF, Saffer DM, Doan ML, et al. (2012) Scale dependence of in situ permeability measurements in the Nankai accretionary prism: The role of fractures. *Geophysical Research Letters* 39: L07302. http://dx.doi.org/10.1029/2012GL051216.

Brace WF (1981) Permeability of crystalline and argillaceous rocks. *International Journal of Rock Mechanics and Mining Science and Geomechanics Abstracts* 17: 241–251.

Brace WF, Walsh JB, and Frango WT (1968) Permeability of granite under high pressure. *Journal of Geophysical Research* 73: 2225–2236.

Bruderer C and Bernabé Y (2001) Network modeling of dispersion: Transition from Taylor dispersion in homogeneous networks to mechanical dispersion in very heterogeneous ones. *Water Resources Research* 37: 897–908.

Bucher K and Grapes R (2011) *Petrogenesis of Metamorphic Rocks*, 8th edn., 427 pp. Heidelberg: Springer.

Burr DT, Sudicky EA, and Naff RL (1994) Nonreactive and reactive solute transport in three-dimensional heterogeneous porous media: Mean displacement, plume spreading, and uncertainty. *Water Resources Research* 30: 791–815.

Burridge R and Keller JB (1981) Poroelastic equations derived from microstructure. *Journal of the Acoustical Society of America* 70: 1140–1146.

Butler JJ (2005) Hydrological methods for estimation of spatial variations in hydraulic conductivity. In: Rubin Y and Hubbard SS (eds.) *Hydrogeophysics*, pp. 23–58. Berlin: Springer.

Cade CA, Evans IJ, and Bryant SL (1994) Analysis of permeability controls: A new approach. *Clay Minerals* 29: 491–501.

Charlaix E, Hulin JP, and Plona TJ (1987) Experimental study of tracer dispersion in sintered glass porous materials of variable compaction. *Physics of Fluids* 30: 1690–1698.

Chen J, Hopmans JW, and Grismer ME (1999) Parameter estimation of two-fluid capillary pressure–saturation and permeability functions. *Advances in Water Resources* 22: 479–493.

Choquette PW and Pray LC (1970) Geologic nomenclature and classification of porosity in sedimentary carbonates. *American Association of Petroleum Geologists Bulletin* 54: 207–250.

Chow TS (1998) Wetting of rough surfaces. *Journal of Physics: Condensed Matter* 10: L445–L451.

Choy TC (1999) *Effective Medium Theory: Principles and Applications*, 182 pp. Oxford: Oxford University Press.

Clavaud J-B, Maineult A, Zamora M, Rasolofosaon P, and Schlitter C (2008) Permeability anisotropy and its relations with porous medium structure. *Journal of Geophysical Research* 113: B01202. http://dx.doi.org/10.1029/2007JB005004.

Clennell MB (1997) Tortuosity: A guide through the maze. *Geological Society of London, Special Publication* 122: 299–344.

Colón CFJ, Oelkers EH, and Schott J (2004) Experimental investigation of the effect of dissolution on sandstone permeability, porosity, and reactive surface area. *Geochimica et Cosmochimica Acta* 68: 805–817.

Cortis A and Berkowitz B (2004) Anomalous transport in "classical" soil and sand columns. *Soil Science Society of America Journal* 68: 1539–1548.

Dagan G (1989) *Flow and Transport in Porous Formations*, 465 pp. Heidelberg: Springer.

David C, Wong T-F, Zhu W, and Zhang J (1994) Laboratory measurement of compaction-induced permeability change in porous rock: Implications for the generation and maintenance of pore pressure excess in the crust. *Pure and Applied Geophysics* 143: 425–456.

de Gennes PG (1983) Theory of slow biphasic flows in porous media. *Physicochemical Hydrodynamics* 4: 175–185.

de Gennes PG (1985) Wetting: Statics and dynamics. *Reviews of Modern Physics* 57: 827–863.

DeHoff RT, Aigeltinger EH, and Craig KR (1972) Experimental determination of the topological properties of three-dimensional microstructures. *Journal of Microscopy* 95: 69–91.

Dewhurst DN, Aplin AC, Sarda J-P, and Yang Y (1998) Compaction-driven evolution of porosity and permeability in natural mudstones: An experimental study. *Journal of Geophysical Research* 103: 651–661.

Doughty DA and Tomutsa L (1997) Imaging pore structure and connectivity in high resolution NMR microscopy. *International Journal of Rock Mechanics and Mining Science* 34: 3–4.

Dullien FAL (1992) *Porous Media: Fluid Transport and Pore Structure*, 2nd edn., 574 pp. New York: Academic Press.

Dunoyer de Segonzac G (1970) The transformation of clay minerals during diagenesis and low-grade metamorphism: A review. *Sedimentology* 15: 281–346.

Ehrenberg SN, Eberli GP, Keramati M, and Moallemi SA (2006) Porosity–permeability relationships in interlayered limestone-dolostone reservoirs. *American Association of Petroleum Geologists Bulletin* 90: 91–114.

Evans B, Bernabé Y, and Zhu W (1999) Evolution of pore structure and permeability of rocks in laboratory experiments. In: Jamviet B and Meakin P (eds.) *Growth, Dissolution and Pattern Formation in Geosystems*, pp. 327–344. Boston, MA: Kluwer.

Farmer CL (2002) Upscaling: A review. *International Journal for Numerical Methods in Fluids* 40: 63–78.

Ferreol B and Rothman D (1995) Lattice-Boltzmann simulations of flow through Fontainebleau sandstone. *Transport in Porous Media* 20: 3–20.

Fischer GJ (1992) The determination of permeability and storage capacity: Pore pressure oscillation method. In: Evans B and Wong T-F (eds.) *Fault Mechanics and Transport Properties of Rocks: A Festschrift in Honor of W.F. Brace*, pp. 187–211. San Diego, CA: Academic Press.

Folk RL (1959) Practical petrographic classification of limestones. *American Association of Petroleum Geologists Bulletin* 43: 1–38.

Fredrich JT (1999) 3D imaging of porous media using laser scanning confocal microscopy with application to microscale transport processes. *Physics and Chemistry of the Earth, Part A* 24: 551–561.

Fredrich JT, Menéndez B, and Wong T-F (1995) Imaging pore structure of geomaterials. *Science* 268: 276–278.

Gaillard J-F, Chen C, Stonedahl SH, Lau BLT, Keane DT, and Packman AI (2007) Imaging of colloidal deposits in granular porous media by X-ray difference micro-tomography. *Geophysical Research Letters* 34: L18404. http://dx.doi.org/10.1029/2007GL030514.

Gamage K, Screaton E, Bekins B, and Aiello I (2011) Permeability–porosity relationships of subduction zone sediments. *Marine Geology* 279: 19–36.

Gangi AF (1978) Variation of whole and fractured porous rock permeability with confining pressure. *International Journal of Rock Mechanics and Mining Science* 15: 249–257.

Gavrilenko P and Guéguen Y (1989) Pressure dependence of permeability: A model for cracked rocks. *Geophysical Journal International* 98: 159–172.

Gelhar LW (1993) *Stochastic Subsurface Hydrology*. Englewood Cliffs, 390 pp. NJ: Prentice Hall.
Gelhar LW, Welty C, and Rehfeldt KR (1992) A critical review of data on field-scale dispersion in aquifers. *Water Resources Research* 28: 1955–1974.
Ghabezloo S, Sulem J, and Saint-Marc J (2009) Evaluation of a permeability–porosity relationship in slow-permeability creeping material using a single transient test. *International Journal of Rock Mechanics and Mining Science* 46: 761–768.
Gouze P, Melean Y, Le Borgne T, Dentz M, and Carrera J (2008) Non-Fickian dispersion in porous media explained by heterogeneous microscale matrix diffusion. *Water Resources Research* 44: W11416. http://dx.doi.org/10.1029/2007WR006690.
Gray WG and O'Neill K (1976) On the general equations for flow in porous media and their reduction to Darcy's law. *Water Resources Research* 12: 148–154.
Gribbin J, Zhu W, and Tivey MK (2012) Anisotropy in seafloor flange, slab, and crust samples from measurements of permeability and porosity: Implications for fluid flow and deposit evolution. *Geochemistry, Geophysics, Geosystems* 13: Q03018. http://dx.doi.org/10.1029/2011GC003840.
Gudmundsson A (2011) *Rock Fractures in Geological Processes*, 578 pp. Cambridge: Cambridge University Press.
Guéguen Y and Dienes J (1989) Transport properties of rocks from statistics and percolation. *Mathematical Geology* 21: 1–13.
Guéguen Y, Gavrilenko P, and Le Ravalec M (1996) Scales of rock permeability. *Surveys in Geophysics* 17: 245–263.
Guéguen Y and Palciauskas V (1994) *Introduction to the Physics of Rocks*, 297 pp. Princeton, NJ: Princeton University Press.
Guin JA, Kessler DP, and Greenkorn RA (1972) The dispersion tensor in anisotropic porous media. *Industrial and Engineering Chemistry Fundamentals* 11: 477–482.
Gunstensen AK and Rothman DH (1993) Lattice-Boltzmann studies of two-phase flow through porous media. *Journal of Geophysical Research* 98: 6431–6441.
Hristopulos DT (2003) Renormalization group methods in subsurface hydrology: Overview and application in hydraulic conductivity upscaling. *Advances in Water Resources* 26: 1279–1308.
Hunt AG and Ewing R (2009) *Percolation Theory for Flow in Porous Media*, 2nd edn., 319 pp. Heidelberg: Springer.
Jiang Z, Wu K, Couples GS, van Dijke MIJ, Sorbie KS, and Ma J (2007) Efficient extraction of networks from three-dimensional porous media. *Water Resources Research* 43: W12S03. http://dx.doi.org/10.1029/2006WR005780.
Johnson DL (1989) Scaling function for dynamic permeability in porous media. *Physical Review Letters* 63: 580.
Johnson DL, Koplik J, and Dashen R (1987) Theory of dynamic permeability and tortuosity in fluid-saturated porous media. *Journal of Fluid Mechanics* 176: 379–402.
Johnson DL, Plona TJ, Scala C, Pasierb F, and Kojima H (1982) Tortuosity and acoustic slow waves. *Physical Review Letters* 49: 1840–1844.
Kamath J, Boyer RE, and Nakagawa FN (1992) Characterization of core-scale heterogeneities using laboratory transients. *SPE Formation Evaluation* 7: 219–227.
Kameda A, Dvorkin J, Keehm Y, Nur A, and Bosl W (2006) Permeability–porosity transforms from small sandstone fragments. *Geophysical Journal of the Royal Astronomical Society* 71: N11–N19.
Keehm Y, Mukerji T, and Nur A (2001) Computational rock physics at the pore scale: Transport properties and diagenesis in realistic pore geometries. *The Leading Edge* 20: 180–183.
King PR (1989) The use of renormalization for calculating effective permeability. *Transport in Porous Media* 4: 37–58.
Kirkpatick S (1973) Percolation and conduction. *Reviews in Modern Physics* 45: 574–588.
Ko SC, Olgaard DL, and Wong TF (1997) Generation and maintenance of pore pressure excess in a dehydrating system: 1, Experimental and microstructural observations. *Journal of Geophysical Research* 102: 195–214.
Levitz P (2007) Toolbox for 3D imaging and modeling of porous media: Relationship with transport properties. *Cement and Concrete Research* 37: 351–359.
Levy M and Berkowitz B (2003) Measurement and analysis of non-Fickian dispersion in heterogeneous porous media. *Journal of Contaminant Hydrology* 64: 203–226.
Li K and Horne RN (2006) Comparison of methods to calculate relative permeability from capillary pressure in consolidated water-wet porous media. *Water Resources Research* 42: W06405. http://dx.doi.org/10.1029/2005WR004482.
Lindquist WB, Lee SM, Coker DA, Jones KW, and Spanne P (1996) Medial axis analysis of three-dimensional tomographic images of drill core samples. *Journal of Geophysical Research* 101: 8297–8310.
Lindquist WB, Venkatarangan A, Dunsmuir J, and Wong T-F (2000) Pore and throat size distribution measured from synchrotron X-ray tomographic images of Fontainebleau sandstone. *Journal of Geophysical Research* 105: 21508–21528.
Lockner DA, Byerlee JD, Kuksenko V, Ponomarev A, and Sidorin A (1991) Quasi-static fault growth and shear fracture energy in granite. *Nature* 350: 39–42.
Luchnikov VA, Medvedev NN, Oger L, and Troadec JP (1999) Voronoi-Delaunay analysis of voids in systems of nonspherical particles. *Physical Review E* 59: 7205–7212.
Luquot L and Gouze P (2009) Experimental determination of porosity and permeability changes induced by injection of CO_2 in carbonate rocks. *Chemical Geology* 265: 148–159.
MacDonald I, Kaufmann P, and Dullien FAL (1986) Quantitative image analysis of finite porous media: 1. Development of genus and pore map software. *Journal of Microscopy* 144: 277–296.
Madras G, Hamilton BL, and Matthews MA (1996) Influence of adsorption on the measurement of diffusion coefficients by Taylor dispersion. *International Journal of Thermophysics* 17: 373–389.
Maineult A, Strobach E, and Renner J (2008) Self-potential signals induced by periodic pumping tests. *Journal of Geophysical Research* 113: B01203. http://dx.doi.org/10.1029/2007JB005193.
Mallon AJ, Swarbrick RE, and Katsube TJ (2005) Permeability of fine-grained rocks: New evidence from chalks. *Geology* 33: 21–24.
Manz B, Gladden LF, and Warren PB (1999) Flow and dispersion in porous media: Lattice-Boltzmann and NMR studies. *AIChE Journal* 45: 1845–1854.
Mavko G, Mukerji T, and Dvorkin J (2003) *The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media*, 331 pp. Cambridge: Cambridge University Press.
Mavko G and Nur A (1997) The effect of a percolation threshold in the Kozeny–Carman relation. *Geophysics* 62: 1480–1482.
McCune CC, Fogler HS, and Kline WS (1979) An experimental technique for obtaining permeability–porosity relationships in acidized porous media. *Industrial and Engineering Chemistry Fundamentals* 18: 188–191.
Medvedev NN, Voloshin VP, Luchnikov VA, and Gavrilova ML (2006) An algorithm for three-dimensional Voronoi S-network. *Journal of Computational Chemistry* 27: 1676–1692.
Mei CC and Auriault JL (1989) Mechanics of heterogeneous porous media with several spatial scales. *Proceedings of the Royal Society of London A* 426: 391–423.
Menéndez B, David C, and Martinez Nistal A (2001) Confocal scanning laser microscopy applied to the study of pore and crack networks in rocks. *Computers and Geosciences* 27: 1101–1109.
Mitchell JK and Soga K (2005) *Fundamentals of Soil Behavior*, 3rd edn., 560 pp. New York: Wiley.
Mok U, Bernabé Y, and Evans B (2002) Permeability, porosity and pore geometry of chemically altered porous silica glass. *Journal of Geophysical Research* 107: B12015. http://dx.doi.org/10.1029/2001JB000247.
Morse JW (2003) Formation and diagenesis of carbonate sediments. *Treatise on Geochemistry*, vol. 7, pp. 67–85. Amsterdam: Elsevier.
Mueller S, Melnik O, Spieler O, Scheu B, and Dingwell DB (2005) Permeability and degassing of dome lavas undergoing rapid decompression: An experimental determination. *Bulletin of Volcanology* 67: 526–538.
Naff RL (1990) On the nature of the dispersive flux in saturated heterogeneous porous media. *Water Resources Research* 26: 1013–1026.
Nelson PH (2000) Evolution of permeability–porosity trends in sandstones. In: *41st Annual Logging Symposium, MM1-14*, SPWLA.
Neuman SP (1977) Theoretical derivation of Darcy's law. *Acta Mechanica* 25: 153–170.
Noetinger B, Estebenet T, and Quintard M (2001) Up-scaling flow in fractured media: Equivalence between the large scale averaging theory and the continuous time random walk. *Transport in Porous Media* 43: 581–596.
Øren P-E and Bakke S (2002) Process based reconstruction of sandstones and prediction of transport properties. *Transport in Porous Media* 46: 311–343.
Panfilov M (2000) *Macroscale Models of Flow Through Highly Heterogeneous Porous Media*, 367 pp. Dordrecht: Kluwer Academic Press.
Papafotiou A, Helmig R, Schaap J, et al. (2008) From the pore scale to the lab scale: 3-D lab experiment and numerical simulation of drainage in heterogeneous porous media. *Advances in Water Resources* 31: 1253–1268.
Pape H, Clauser C, and Iffland J (2000) Variation of permeability with porosity in sandstone diagenesis interpreted with a fractal pore space model. *Pure and Applied Geophysics* 157: 603–619.
Passchier CW and Trouw RAJ (2005) *Microtectonics*, 2nd edn., 371 pp. Heidelberg: Springer.
Paterson MS (1983) The equivalent channel model for permeability and resistivity in fluid-saturated rock – A re-appraisal. *Mechanics of Materials* 2: 345–352.
Persoff P and Pruess K (1995) Two-phase flow visualization and relative permeability measurement in natural rough-walled rock fractures. *Water Resources Research* 31: 1175–1186.
Petford N, Davidson G, and Miller JA (1999) Pore structure determination using confocal scanning laser microscopy. *Physics and Chemistry of the Earth, Part A* 24: 563–567.

Petford N, Davidson G, and Miller JA (2001) Investigation of the petrophysical properties of a porous sandstone sample using confocal scanning laser microscopy. *Petroleum Geoscience* 7: 99–105.

Pettijohn FJ (1975) *Sedimentary Rocks*, 3rd edn., 628 pp. New York: Harper and Row.

Pettijohn FJ, Potter PE, and Siever R (1987) *Sand and Sandstone*, 2nd edn., 553 pp. Heidelberg: Springer.

Pittman ED (1979) Recent advances in sandstone diagenesis. *Annual Review of Earth and Planetary Sciences* 7: 39–62.

Poutet J, Manzoni D, Hage-Chehade F, et al. (1996) The effective mechanical properties of reconstructed porous media. *International Journal of Rock Mechanics and Mining Science and Geomechanics Abstracts* 33: 409–415.

Pride SR (1994) Governing equations for the coupled electromagnetics and acoustics of porous media. *Physical Review B* 50: 15678–15696.

Pride SR, Gangi AF, and Morgan FD (1992) Deriving the equations of motion for porous isotropic media. *Journal of the Acoustical Society of America* 92: 3278–3290.

Ramsay JG (1967) *Folding and Fracturing of Rocks*, 568 pp. New York: McGraw-Hill.

Reis JC and Acock AM (1994) Permeability reduction models for the precipitation of inorganic solids in Berea sandstone. *Journal of Geophysical Research* 18: 347–368.

Renard P and de Marsily G (1997) Calculating equivalent permeability: A review. *Advances in Water Resources* 20: 253–278.

Renard P, Genty A, and Stauffer F (2001) Laboratory determination of the full permeability tensor. *Journal of Geophysical Research* 106: 26443–26452.

Revil A, Cary L, Fan Q, Finizola A, and Trolard F (2005) Self-potential signals associated with preferential ground water flow pathways in a buried paleo-channel. *Geophysical Research Letters* 32: L07401. http://dx.doi.org/10.1029/2004GL022124.

Revil A and Cathles LM (1999) Permeability of shaly sands. *Water Resources Research* 35: 651–662.

Revil A and Glover PWJ (1997) Theory of ionic-surface electrical conduction in porous media. *Physical Review B* 55: 1757–1773.

Revil A and Glover PWJ (1998) Nature of surface electrical conductivity in natural sands, sandstones and clays. *Geophysical Research Letters* 25: 691–694.

Revil A and Leroy P (2004) Constitutive equations for ionic transport in porous shales. *Journal of Geophysical Research* 109: B03208. http://dx.doi.org/10.1029/2003JB002755.

Revil A and Linde N (2006) Chemico-electromechanical coupling in microporous media. *Journal of Colloid and Interface Science* 302: 682–694.

Revil A and Pessel M (2002) Electroosmotic flow and the validity of the classical Darcy equation in silty shales. *Geophysical Research Letters* 29: 1300. http://dx.doi.org/10.1029/2001GL013480.

Rizzo E, Suski B, Revil A, Straface S, and Troisi S (2004) Self-potential signals associated with pumping tests experiments. *Journal of Geophysical Research* 109: B10203. http://dx.doi.org/10.1029/2004JB003049.

Rothman DH (1990) Macroscopic laws for immiscible two-phase flow in porous media: Results from numerical experiments. *Journal of Geophysical Research* 95: 8663–8674.

Rubin Y and Hubbard SS (2005) *Hydrogeophysics*, 523 pp. Berlin: Springer.

Saar MO and Manga M (1999) Permeability–porosity relationship in vesicular basalts. *Geophysical Research Letters* 26: 111–114.

Sahimi M (1995) *Flow and Transport in Porous Media and Fractured Rock*. Weinheim: VCH, 482 pp.

Sahimi M, Hughes BD, Scriven LE, and Davis HT (1986) Dispersion in flow through porous media – I. One-phase flow. *Chemical Engineering Science* 41: 2103–2122.

Sallès J, Thovert JF, and Adler PM (1993) Reconstructed porous media and their application to fluid flow and solute transport. *Journal of Contaminant Hydrology* 13: 3–22.

Scott GJT, Webster R, and Nortcliff S (1988) The topology of pore structure in cracking clay soil: II. Connectivity density and its estimation. *Journal of Soil Science* 39: 315–326.

Silin DB, Jin G, and Patzek TW (2004) Robust determination of the pore space morphology in sedimentary rocks. *Journal of Petroleum Technology* 56: 69–70.

Slattery JC (1967) Flow of viscoelastic fluids through porous media. *AIChE Journal* 13: 1066–1071.

Sok RM, Knackstedt MA, Sheppard AP, et al. (2002) Direct and stochastic generation of network models from tomographic images: Effect of topology on residual saturations. *Transport in Porous Media* 46: 345–372.

Song I, Elphick SC, Main IG, Ngwenya BT, Odling NW, and Smyth NF (2004) One-dimensional fluid diffusion induced by constant-rate flow injection: Theoretical analysis and application to the determination of fluid permeability and specific storage of a cored rock sample. *Journal of Geophysical Research* 109: B05207. http://dx.doi.org/10.1029/2003JB002395.

Song I and Renner J (2006) Linear pressurization method for determining hydraulic permeability and specific storage of a rock sample. *Geophysical Journal International* 164: 685–696.

Spangenberg E, Spangenberg U, and Heindorf C (1998) An experimental study of transport properties of porous rock salt. *Physics and Chemistry of the Earth* 23: 367–371.

Spanne P, Thovert JF, Jacquin CJ, Lindquist WB, Jones KW, and Adler PM (1994) Synchrotron computed microtomography of porous media: Topology and transport. *Physical Review Letters* 73: 2001–2004.

Stanchits S, Vinciguerra S, and Dresen G (2006) Ultrasonic velocities, acoustic emission characteristics and crack damage of basalt and granite. *Pure and Applied Geophysics* 163: 974–993.

Tallakstad KT, Knudsen HA, Ramstad T, et al. (2009) Steady-state two-phase flow in porous media: Statistics and transport properties. *Physical Review Letters* 102: 074502, 4 pages. http://dx.doi.org/10.1103/PhysRevLett.102.074502.

Tenthorey E and Cox SF (2003) Reaction-enhanced permeability during serpentinite dehydration. *Geology* 31: 921–924.

Underwood EE (1970) *Quantitative Stereology*, 274 pp. Reading, MA: Addison Wesley.

Van Den Broeck C (1982) Stochastic description of longitudinal dispersion in uniaxial flows. *Physica A* 112: 343–352.

Van Den Broeck C (1990) Taylor dispersion revisited. *Physica A* 168: 677–696.

Vereecken H, Kasteel R, Vanderborght J, and Harter T (2007) Upscaling hydraulic properties and soil water flow processes in heterogeneous soils: A review. *Vadose Zone Journal* 6: 1–28.

Vogel HJ (1997) Morphological determination of pore connectivity as a function of pore size using serial sections. *European Journal of Soil Science* 48: 365–377.

Vogel HJ (2002) Topological characterization of Porous Media. In: Mecke KR and Stoyan D (eds.) *Morphology of Condensed Matter*, pp. 75–92. Heidelberg: Springer.

Vogel HJ and Roth K (2001) Quantitative morphology and network representation of soil pore structure. *Advances in Water Resources* 24: 233–242.

Walsh JB (1981) Effect of pore pressure and confining pressure on fracture permeability. *International Journal of Rock Mechanics and Mining Science* 18: 429–435.

Walsh JB and Brace WF (1984) Effect of pressure on porosity and the transport properties of rock. *Journal of Geophysical Research* 89: 9425–9431.

Wark DA and Watson EB (1998) Grain-scale permeabilities of texturally equilibrated monomineralic rocks. *Earth and Planetary Science Letters* 164: 591–605.

Whitaker S (1986a) Flow in porous media, I: A theoretical derivation of Darcy's law. *Transport in Porous Media* 1: 3–25.

Whitaker S (1986b) Flow in porous media, II: The governing equations for immiscible, two-phase flow. *Transport in Porous Media* 1: 105–125.

Whitaker S (1999) *The Method of Volume Averaging*, 223 pp. Dordrecht: Kluwer Academic Press.

Wildenschild D, Hopmans JW, Vaz CMP, Rivers ML, Rikard D, and Christensen BSB (2002) Using X-ray computed tomography in hydrology: Systems, resolutions and limitations. *Journal of Hydrology* 267: 285–297.

Wilson MJ (1999) The origin and formation of clay minerals in soils: Past, present and future perspectives. *Clay Minerals* 34: 7–25.

Xiao X, Evans B, and Bernabé Y (2006) Permeability evolution during non-linear viscous creep of calcite rocks. *Pure and Applied Geophysics* 163: 2071–2102.

Yang Y and Aplin AC (2010) A permeability–porosity relationship for mudstones. *Marine and Petroleum Geology* 27: 1692–1697.

Zhang S, Cox SF, and Paterson MS (1994) The influence of room temperature deformation on porosity and permeability in calcite aggregates. *Journal of Geophysical Research* 99: 15761–15775.

Zhang S, Paterson MS, and Cox SF (2001) Microcrack growth and healing in deformed calcite aggregates. *Tectonophysics* 335: 17–36.

Zhu W, Gaetani G, Fusseis F, Montési L, and De Carlo F (2011) Microtomography of partially molten rocks: Three-dimensional melt distribution in mantle peridotite. *Science* 332(6025): 88–91. http://dx.doi.org/10.1126/science.120222.

Zhu W, Montési L, and Wong T-F (2002) Influence of loading path on anisotropic damage and permeability evolution during mechanical compaction of sandstone. In: de Meer S, Drury MR, de Bresser JHP, and Pennock GM (eds.) *Deformation Mechanisms, Rheology and Tectonics: Current Status and Future Perspectives*, *Geological Society, London, Special Publications*, vol. 200, 119–136.

Zhu W, Montési L, and Wong T-F (2007a) A probabilistic damage model of stress-induced permeability anisotropy during cataclastic flow. *Journal of Geophysical Research* 112: B10207. http://dx.doi.org/10.1029/2006JB004456.

Zhu W, Tivey MK, Gittings H, and Craddock PR (2007b) Permeability–porosity relationships in seafloor vent deposits: Dependence on pore evolution processes. *Journal of Geophysical Research* 112: B05208. http://dx.doi.org/10.1029/2006JB004716.

Zhu W and Wong T-F (1997) The transition from brittle faulting to cataclastic flow: Permeability evolution. *Journal of Geophysical Research* 102: 3027–3041.

Zinszner B and Pellerin EM (2007) *A Geoscientist's Guide to Petrophysics*, 384 pp. Paris: Technip.

Zuo L, Krevor S, Falta RW, and Benson SM (2011) An experimental study of CO_2 exsolution and relative permeability measurements during CO_2 saturated water depressurization. *Transport in Porous Media* 91: 459–478. http://dx.doi.org/10.1007/s11242-011-9854-2.

This page intentionally left blank

11.03 Geophysical Properties of the Near Surface Earth: Seismic Properties

DR Schmitt, University of Alberta, Edmonton, AB, Canada

© 2015 Elsevier B.V. All rights reserved.

11.03.1	**Introduction**	44
11.03.2	**Basic Theory**	45
11.03.2.1	Hooke's Constitutive Relationship and Moduli	45
11.03.3	**Mineral Building Blocks**	47
11.03.3.1	Elastic Properties of Minerals	47
11.03.3.2	Bounds on Isotropic Mixtures of Anisotropic Minerals	48
11.03.3.3	Isothermal Versus Adiabatic Moduli	49
11.03.3.4	Effects of Pressure and Temperature on Mineral Moduli	51
11.03.3.5	Mineral Densities	51
11.03.4	**Fluid Properties**	53
11.03.4.1	Phase Relations for Fluids	54
11.03.4.2	Equations of State for Fluids	54
11.03.4.2.1	Ideal gas law	56
11.03.4.2.2	Adiabatic and isothermal fluid moduli	56
11.03.4.2.3	The van der Waals model	57
11.03.4.2.4	The Peng–Robinson EOS	57
11.03.4.2.5	Correlative EOS models	58
11.03.4.2.6	Determining K_f from equations of state	58
11.03.4.3	Mixtures and Solutions	59
11.03.4.3.1	Frozen mixtures	59
11.03.4.3.2	Miscible fluid mixtures	60
11.03.5	**The Rock Frame**	65
11.03.5.1	Essential Characteristics	65
11.03.5.2	The Pore-Free Solid Portion	66
11.03.5.3	Influence of Porosity	67
11.03.5.4	Influence of Crack-Like Porosity	69
11.03.5.5	Pressure Dependence in Granular Materials	72
11.03.5.6	Implications of Pressure Dependence	73
11.03.5.6.1	Stress-induced anisotropy (acoustoelastic effect)	73
11.03.5.6.2	Influence of pore pressure	74
11.03.6	**Seismic Waves in Fluid-Saturated Rocks**	74
11.03.6.1	Gassmann's Equation	75
11.03.6.2	Frequency-Dependent Models	76
11.03.6.2.1	Global flow (biot) model	77
11.03.6.2.2	Local flow (squirt) models	78
11.03.7	**Empirical Relations and Data Compilations**	78
11.03.8	**The Road Ahead**	81
Acknowledgments		81
References		81

Glossary

Adiabat An adiabatic path in P–V–T space.

Adiabatic A thermodynamic process in which no heat is allowed to transfer into or out of the system. The local compression and rarefaction and corresponding increase and decrease of both pressure and temperature of a material as compressional wave passes are assumed to be an adiabatic process.

Anisotropy The condition in which the physical properties of a material will depend on direction.

Aspect ratio χ (dimensionless) In the context of crack-like porosity, this refers to the aperture width of the crack to its length.

Bulk modulus K (Pa) Also called the incompressibility. A measure of the resistance of a material to deformation for a given change in pressure.

Compliances S_{ij} (Pa^{-1}) The elastic mechanical parameters that generally relate stresses to strains.

Compressibility (Pa^{-1}) Inverse of the bulk modulus.

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00190-1

Cricondenbar For fluid mixtures. The greatest pressure at which both liquid and vapor phases can coexist. Above the cricondenbar, the mixture must be either a liquid or a supercritical fluid phase.
Critical point For pure fluids, a point in P–V–T space at which the liquid–vapor phase line terminates. The fluid will be in the supercritical state for pressures and temperatures above the critical pressure P_c and temperature T_c. At the critical point the fluid will have the critical specific volume V_c or equivalently the critical density $\rho_c = M/V_c$, where M is the chemical molecular weight.
Cricondentherm For fluid mixtures. The greatest temperature at which both liquid phase and vapor phase can still coexist. Above this temperature, the fluid will be either vapor or supercritical fluid phase.
Density ρ (kg m^{-3}) Mass per unit volume.
Equation of state A theoretical or empirical function or set of functions that describes the material's specific volume as a function of pressure and temperature.
Hooke's law The mathematical relationship between stress and strain via the elastic stiffnesses or conversely the strains and the stresses via the elastic compliances.
Isentropic A thermodynamic process in which the entropy of the system remains constant. A reversible adiabatic process is also isentropic.
Isochor A thermodynamic path in P–V–T space in which the specific volume V_m or the density ρ remains constant.
Isotherm A thermodynamic path in P–V–T space in which the temperature T remains constant. These are often the conditions employed in conventional measurements of fluid properties particularly in the petroleum industry.
Lamé parameters λ and μ (Pa). The two elastic parameters relating stresses to strains in the Lamé mathematical formulation of Hooke's law.
Poisson's ratio ν (dimensionless) The negative of the ratio between the radial and the axial strains induced by an axial stress.
Polycrystal A material that is a mixture of mineral crystals and that, often, is assumed to be free of pores. The properties of the polycrystal are then taken to be representative of those for the solid portion of the rock.
Pseudocritical point For fluid mixtures, a point in P–V–T space where the bubble and dew lines meet. This point depends on the composition of the mixture and occurs at the pseudocritical pressure P_{PC} and temperature T_{PC}.
Saturated The condition where the pore space of the rock is filled with fluids.
Saturation The fraction of the pore space that is filled with a given fluid. If only one fluid fills the pore volume, it will have a saturation of 1. If the pore volume is equally filled with two different fluids, they each will have a saturation of 0.5.
Shear modulus μ (Pa) The elastic mechanical parameter relating shear stress to shear strain.
Stiffness C_{ij} (Pa) The elastic mechanical parameters that generally relate strains to stresses.
Strain ϵ_{ij} or γ_{ij} (dimensionless) Measures of the deformation of a material.
Stress σ_{ij} or τ_{ij} (Pa) The ratio of an applied force to the area over which it is applied. Normal stresses σ_{ij} are directed perpendicularly to the surface. Shear stresses τ_{ij} are directed along the plane of the surface.
Supercritical The condition for a fluid encountered in P–V–T space in which it is no longer considered a liquid or a vapor (gas) but a fluid with the characteristics of both. For single-component fluids, the supercritical phase exists above the critical point at the critical pressure P_c and temperature T_c.
Young's modulus E (Pa) Also often referred to as the modulus of elasticity. The elastic mechanical parameter relating the linear axial strain induced to the applied axial normal stress.

11.03.1 Introduction

Geophysicists measure the spatial and temporal variations in electromagnetic, magnetic, and gravitational potentials and seismic wave fields in order to make inferences regarding the internal structure of the Earth in terms of, respectively, its electrical resistivity (See Chapters 2.25, 11.04, 11.08, and 11.10), its magnetism (See Chapters 2.24, 5.08, 11.05, 11.11), its density (See Chapter 3.03, 11.05, 11.12), and its elasticity (See Chapters 1.26 and 2.12). In seismology the most basic observation is that of a seismic wave's travel time from its source to the point of measurement. Seismologists continue to develop increasingly sophisticated analyses to convert this basic observation into seismic velocities from which the Earth's structure may be deduced. This holds true for the simplest 1-D seismic refraction analysis to the most complicated modern 3-D whole Earth tomogram. The influence of this velocity is not so directly apparent in seismic reflection profiles, but proper imaging depends critically on solid knowledge of the *in situ* seismic velocity structure. Indeed, as computational power grows, the differences between inversion and advanced prestack migration in imaging will become less distinct.

Velocity, as it is used in the geophysical community for wave speed, would certainly first come to a geophysicist's mind as a seismic property. It is also the seismic property that is most often used to infer lithology. Liberally, compressional wave velocities that can exist in crustal materials can range from a few hundred meters per second in air-saturated unconsolidated sediments to upward of 8 km s^{-1} for high-grade metamorphosed rocks at the top of the mantle. Typically then, within a given geologic context, the velocities themselves or additional parameters derived from them such as the compressional/shear wave speed ratio V_P/V_S, Poisson's ratio ν, or the seismic parameter $\varphi = V_P^2 - 4V_S^2/3$ are useful indicators of lithology. Unfortunately, the seismic velocities of any given lithology are not unique. Seismic velocities are affected by numerous factors such as mineralogical composition, texture, porosity, fluids, confining stress, pressure, and temperature, all of which contribute by differing degrees to the value of the

observed wave speed. Seismic anisotropy, being the variation of the wave speeds with direction of propagation through the medium, may also need to be included (Chapter 2.20). With attenuation, even the frequency at which the observation is made should be considered.

With a particular focus on the porous and fluid-filled Earth materials near the Earth's surface, the purpose of this contribution is to review 'seismic properties.' To most workers, this will again mean the measurable 'seismic velocity.' Such velocities are what we, as remote observers, can measure. More fundamentally, however, the seismic velocities are a manifestation of the competition between a material's internal forces (represented in a continuum via the elastic moduli) and inertia (through density). Without derivation, the relationships between the compressional P and S waves and the material's moduli and bulk density ρ are

$$V_P = \sqrt{\frac{K + 4\mu/3}{\rho}} \quad [1]$$

$$V_S = \sqrt{\frac{\mu}{\rho}} \quad [2]$$

where K and μ are the bulk and shear moduli for an isotropic medium, respectively. As fluids cannot support a shear stress, their $\mu = 0$ and eqn [1] reduces to the simpler longitudinal elastic wave of speed V_L

$$V_L = \sqrt{\frac{K}{\rho}} \quad [3]$$

first derived by Newton and thermodynamically corrected by Laplace. Care must be exercised in the choice of K particularly for fluids; much of the task of this contribution will not be in the examination of the wave speeds so much as attempting to understand the material's elastic moduli.

Later, when discussing a fluid-filled porous rock, the K and μ to be used in eqns [1] and [2] will necessarily be those of the bulk mixture of solids and fluids and will appropriately be denoted K_{sat} and μ_{sat}, respectively.

A given wave speed depends on the ratio between the moduli and density. One cannot understand the meaning of an observed seismic velocity, nor calculate it, without sufficient knowledge of these underlying moduli and density. Not surprisingly, then, the complexity of the physics needed to describe a wave velocity increases with the number of the material's characteristics considered. In particular, the introduction to the problem of porosity and mobile fluids multiplies the number of free parameters that can influence the material's moduli and density. Properly describing a near-surface material is substantially more difficult than trying to predict variations in, for example, the Earth's mantle where excursions of only a few percent are considered large!

This review is also carried out from the perspective of rock physics, the factors affecting wave speeds and the models used to predict them are presented. This is all done from the viewpoint of an experimentalist concerned with attempting to test such theories. That said, it must be recognized that often, one cannot have all of the information necessary, nor are the models sufficiently sophisticated, to properly predict a given velocity. This will more often than not be the case in seismic investigations of the near surface for resource exploration or engineering and environmental characterization. As such, at the end, a number of empirical relationships are also described. The basics of elasticity, anisotropy, and poroelasticity are briefly covered in order to set the stage for understanding how wave speeds relate to moduli. In attempting to understand the seismic properties of a fluid-saturated crustal rock, one must first have some understanding of the physical properties of the rock's constituent solid minerals, its saturating fluids, and finally its 'frame.' Minerals and fluids are the basic components in rocks in the upper crust and their behavior first must be studied. This is followed by a review of the factors affecting the rock's frame properties. Finally, these different components affecting rock properties are then integrated using various theories to arrive at estimates of the seismic properties. However, one may not always have available sufficient information to allow for the calculation of the physical properties, and for this case, a number of empirical relations and references to published compilations of observed results are provided. Lack of space restricts delving in detail into the range of issues related to seismic properties, so I conclude with some thoughts about the topics that will be important in the coming decade. A short glossary of terms that are not normally encountered in the geophysical literature is also provided.

11.03.2 Basic Theory

Any understanding of the propagation of mechanical waves rests on basic elasticity theory. The number of texts on this topic is large and little is to be gained here by repeating the basic concepts of stress and strain, the development of Hooke's law, or the construction of the wave equations linking elasticity to wave velocities. I assume the reader will have some basic understanding of elasticity. Some recommended texts covering the basic governing equations at different levels of sophistication include Bower (2010), Fung (1965), and Tadmor et al. (2012). Stein and Wysession (2002) gave a particularly cogent exposition of both isotropic elasticity and the solution to the wave equation particularly as it relates to seismology. Auld (1990) provided an excellent advanced overview of elasticity with good emphasis on elastic anisotropy; the notation styles employed here follow largely from this text. For understanding of more complicated materials, the essentials of poroelasticity can be found in Wang (2000), Bourbié et al. (1987), and Gueguen and Bouteca (2004); of anelasticity in Lakes (2009) and Carcione (2007); and of hyperelasticity (nonlinear elasticity) in Holzapfel (2000). These texts will well cover the details, and only the necessary definitions of moduli within the Voigt representation of Hooke's law are provided.

11.03.2.1 Hooke's Constitutive Relationship and Moduli

In this contribution, we assume all strains are infinitesimal and describe to the first order the material's deformation via the strain tensor ε

$$\epsilon(x,y,z)=\begin{bmatrix}\epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz}\\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz}\\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz}\end{bmatrix}$$

$$=\begin{bmatrix}\dfrac{\partial u_x}{\partial x} & \dfrac{1}{2}\left(\dfrac{\partial u_x}{\partial y}+\dfrac{\partial u_y}{\partial x}\right) & \dfrac{1}{2}\left(\dfrac{\partial u_x}{\partial z}+\dfrac{\partial u_z}{\partial x}\right)\\ \dfrac{1}{2}\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right) & \dfrac{\partial u_y}{\partial y} & \dfrac{1}{2}\left(\dfrac{\partial u_y}{\partial z}+\dfrac{\partial u_z}{\partial y}\right)\\ \dfrac{1}{2}\left(\dfrac{\partial u_z}{\partial x}+\dfrac{\partial u_x}{\partial z}\right) & \dfrac{1}{2}\left(\dfrac{\partial u_z}{\partial y}+\dfrac{\partial u_y}{\partial z}\right) & \dfrac{\partial u_z}{\partial z}\end{bmatrix} \quad [4]$$

where the displacement of the particle originally at (x, y, z) is given by the vector $\mathbf{u}=u_x\mathbf{i}+u_y\mathbf{j}+u_z\mathbf{k}$. Forces within the materials are defined by the stress tensor σ:

$$\sigma(x,y,z)=\begin{bmatrix}\sigma_{xx} & \tau_{xy} & \tau_{xz}\\ \tau_{yx} & \sigma_{yy} & \tau_{yz}\\ \tau_{zx} & \tau_{zy} & \sigma_{zz}\end{bmatrix} \quad [5]$$

where normal and shear stresses are represented by σ_{ii} and τ_{ij}, respectively. Hooke's law is the linear constitutive relationship between strain equation [4] and stress equation [5]. The simplest case of an isotopic medium is most commonly assumed in studies of wave propagation. In this case, Hooke's law may be written in the abbreviated Voigt notation as using the mathematical simplifications afforded by the use of the Lamé parameters λ and μ:

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \tau_{yz}\\ \tau_{zx}\\ \tau_{xy}\end{bmatrix}=\begin{bmatrix}\lambda+2\mu & \lambda & \lambda & 0 & 0 & 0\\ \lambda & \lambda+2\mu & \lambda & 0 & 0 & 0\\ \lambda & \lambda & \lambda+2\mu & 0 & 0 & 0\\ 0 & 0 & 0 & \mu & 0 & 0\\ 0 & 0 & 0 & 0 & \mu & 0\\ 0 & 0 & 0 & 0 & 0 & \mu\end{bmatrix}\begin{bmatrix}\epsilon_{xx}\\ \epsilon_{yy}\\ \epsilon_{zz}\\ 2\epsilon_{yz}\\ 2\epsilon_{zx}\\ 2\epsilon_{xy}\end{bmatrix} \quad [6]$$

The reader should take note of the pattern relating the tensors of eqns [4] and [5] to the stress and strain vectors of the Voigt abbreviated notation of eqn [6] in which the shear strains are multiplied by the awkward factor of 2. The Lamé formulation is mathematically elegant and simple. This simplicity comes with some cost, however, in that λ cannot be directly measured in a simple experiment. In contrast, the second Lamé parameter μ is the shear modulus, which is the simple ratio between an applied shear stress τ and the resulting shear strain $\gamma_{ij}=2\epsilon_{ij}$, a value that can readily be experimentally measured (**Figure 1(c)**).

Two other important moduli in isotropic media are Young's modulus E and the already introduced bulk modulus K. These can be found by conducting simple experiments with clear physical interpretations. E is the ratio between an applied uniaxial stress and its corresponding resulting coaxial linear strain (**Figure 1(b)**). K is the ratio between an applied uniform (hydrostatic) pressure P and the consequent volumetric strain $\theta=\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}$. Poisson's ratio ν is not a modulus but it too is an important and popular measure of a material's deformation under stress. It is simply the negative of the ratio between the lateral ϵ_{xx} and axial ϵ_{yy} strains observed during the same test used to measure E (**Figure 1(b)**).

E, K, λ, μ, or ν can be calculated if any of the other two moduli or parameters are known; extensive conversion tables are readily found (e.g., Birch, 1961; Mavko et al., 2003). Some of these relations are, for convenience,

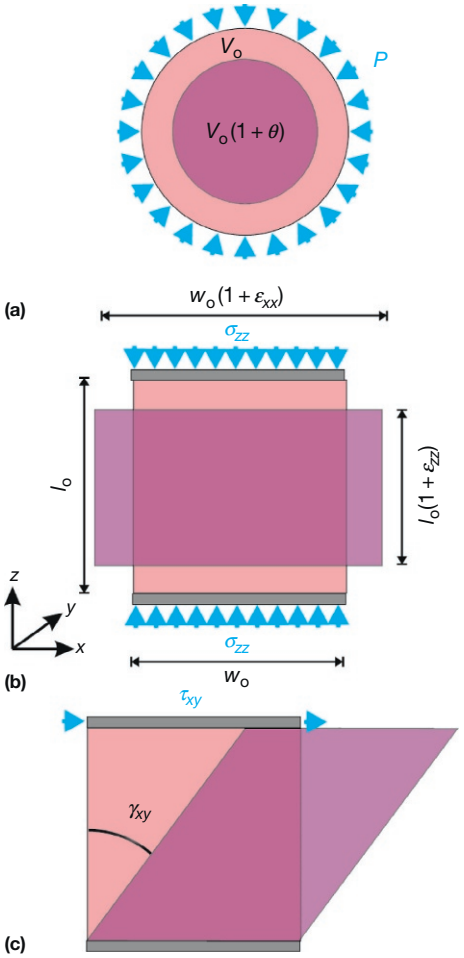


Figure 1 Illustration of the three basic deformations that allow the isotropic elastic moduli to be described. The original dimensions of the object are in light red with the deformed version in transparent purple. (a) Change in volume θV_o upon application of uniform pressure P defining the bulk modulus $K=-P/\theta$, there is no change in shape. (b) Change in the length $l_o\epsilon_{yy}$ and width $w_o\epsilon_{xx}$ upon application of a uniaxial stress σ_{yy} defining the Young's modulus $E=\sigma_{yy}/\epsilon_{yy}$ and Poisson's ratio $\nu=-\epsilon_{xx}/\epsilon_{yy}$. Both the shape and volume change and (c) the change in shape described by the angle $\gamma=2\epsilon_{xy}$ upon application of a simple shear stress τ_{xy} defining the shear modulus $\mu=\tau_{xy}/\epsilon_{xy}$. There is no change in volume.

$$\begin{aligned}E&=\frac{9K\mu}{3K+\mu}\\ K&=\lambda+2\mu/3\\ \mu&=\frac{3}{2}(K-\lambda)\\ \nu&=\frac{3K-E}{6K}\end{aligned} \quad [7]$$

In the case of a liquid, $\mu=0$ and λ can be assigned a physical interpretation as it collapses to the bulk modulus K.

Although we will not be directly addressing issues of seismic reflectivity here, one can also consider the acoustic impedances $Z_i=\rho_i V_i$, where i indicates either the P or the S wave, as a physical property in their own right.

Most of the theoretical models that will follow attempt to develop expressions for K and μ. With this in mind, Hooke's

law (eqn [6]) may alternatively be expressed less elegantly but more physically as

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \tau_{yz}\\ \tau_{zx}\\ \tau_{xy}\end{bmatrix}=\begin{bmatrix}K+\frac{4}{3}\mu & K-\frac{2}{3}\mu & K-\frac{2}{3}\mu & 0 & 0 & 0\\ K-\frac{2}{3}\mu & K+\frac{4}{3}\mu & K-\frac{2}{3}\mu & 0 & 0 & 0\\ K-\frac{2}{3}\mu & K-\frac{2}{3}\mu & K+\frac{4}{3}\mu & 0 & 0 & 0\\ 0 & 0 & 0 & \mu & 0 & 0\\ 0 & 0 & 0 & 0 & \mu & 0\\ 0 & 0 & 0 & 0 & 0 & \mu\end{bmatrix}\begin{bmatrix}\epsilon_{xx}\\ \epsilon_{yy}\\ \epsilon_{zz}\\ 2\epsilon_{yz}\\ 2\epsilon_{zx}\\ 2\epsilon_{xy}\end{bmatrix} \quad [8]$$

A comparison of eqns [1] and [7] also allows the moduli to be written in terms of the wave speeds as

$$\begin{aligned}\mu&=\rho V_S^2,\\ K&=\rho\left(V_P^2-\frac{4}{3}V_S^2\right),\\ \nu&=\frac{1}{2}\left(\frac{V_P^2-2V_S^2}{V_P^2-V_S^2}\right)\end{aligned} \quad [9]$$

11.03.3 Mineral Building Blocks

11.03.3.1 Elastic Properties of Minerals

The majority of seismological studies assume that the elastic responses of Earth materials behave according to eqn [6]. In reality, however, isotropy is the exception; all minerals and most rocks are elastically anisotropic. We introduce this topic early as it is key to understanding how we arrive at the physical properties of the minerals that constitute the rocks. To include this anisotropy, eqn [8] may more generally be rewritten as

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \tau_{yz}\\ \tau_{zx}\\ \tau_{xy}\end{bmatrix}=\sigma=\begin{bmatrix}C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16}\\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26}\\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36}\\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46}\\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56}\\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66}\end{bmatrix}=\begin{bmatrix}\epsilon_{xx}\\ \epsilon_{yy}\\ \epsilon_{zz}\\ 2\epsilon_{yz}\\ 2\epsilon_{zx}\\ 2\epsilon_{xy}\end{bmatrix}=[c]\epsilon \quad [10]$$

where each of the matrix components C_{ij} is called the elastic stiffnesses. More formally, eqn [10] condenses the expression of Hooke's law through the fourth-order tensor components $\sigma_{ij}=c_{ijkl}\epsilon_{kl}$ with the condensed subscripts following the rules:

$$\mathbf{xx}\rightarrow\mathbf{1},\ \mathbf{yy}\rightarrow\mathbf{2},\ \mathbf{zz}\rightarrow\mathbf{3},\ \mathbf{yz}\rightarrow\mathbf{4},\ \mathbf{xz}\rightarrow\mathbf{5},\ \text{and}\ \mathbf{xy}\rightarrow\mathbf{6} \quad [11]$$

and with the full stiffness matrix represented by $[c]$. ϵ and σ represent the corresponding stress and strain vectors, respectively, within this abbreviated notation. Further, $C_{ij}=C_{ji}$ making $[c]$ symmetrical such that there are at most 21 independent elastic stiffnesses. The total number depends on the degree of symmetry of the system varying from only 2 for the isotropic case just shown to the full 21 for the least symmetrical triclinic crystals as will be discussed shortly. Conversely, the strains can be written in terms of the applied stresses using the elastic compliances S_{ij} with their matrix similarly represented by $[s]$

$$\begin{bmatrix}\epsilon_{xx}\\ \epsilon_{yy}\\ \epsilon_{zz}\\ 2\epsilon_{yz}\\ 2\epsilon_{zx}\\ 2\epsilon_{xy}\end{bmatrix}=\epsilon=\begin{bmatrix}S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16}\\ S_{21} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26}\\ S_{31} & S_{32} & S_{33} & S_{34} & S_{35} & S_{36}\\ S_{41} & S_{42} & S_{43} & S_{44} & S_{45} & S_{46}\\ S_{51} & S_{52} & S_{53} & S_{54} & S_{55} & S_{56}\\ S_{61} & S_{62} & S_{63} & S_{64} & S_{65} & S_{66}\end{bmatrix}\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \tau_{yz}\\ \tau_{zx}\\ \tau_{xy}\end{bmatrix}=[s]\sigma \quad [12]$$

The author takes no responsibility for the fact that conventionally $[s]$ and $[c]$ are used to, respectively, denote the compliances and the stiffnesses. Regardless, the Voigt form is particularly useful because

$$[c]=[s]^{-1} \quad [13]$$

where the superscript '-1' indicates the matrix inverse operation (Auld, 1990). We note that an important property of the tensor (in its abbreviated form here) is its transformation upon change of the basis coordinate system. The mathematics describing this is beyond the needs of this contribution but Walker and Wookey (2012) provided convenient MATLAB® codes to carry out these otherwise tedious calculations.

Using this form, eqn [8] for the isotropic case is

$$\begin{bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \tau_{yz}\\ \tau_{zx}\\ \tau_{xy}\end{bmatrix}=\begin{bmatrix}C_{11} & C_{12} & C_{12} & 0 & 0 & 0\\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0\\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0\\ 0 & 0 & 0 & C_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & C_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & C_{44}\end{bmatrix}\begin{bmatrix}\epsilon_{xx}\\ \epsilon_{yy}\\ \epsilon_{zz}\\ 2\epsilon_{yz}\\ 2\epsilon_{zx}\\ 2\epsilon_{xy}\end{bmatrix} \quad [14]$$

with only C_{11} and C_{12} necessary and $C_{44}=(C_{11}-C_{12})/2$. The symmetry of $[c]$ across the main diagonal should be kept in mind.

Some of the hypothetical experiments that could be carried out in order to find some of the rarer off-axis elastic stiffnesses that couple normal stresses to shear strains and vice versa are illustrated in **Figure 2**. Hearmon (1946) gave the geometry for a number of such static measurements that could be used to fill $[c]$.

For anisotropic materials, the configuration of $[c]$ depends on the material's degree of symmetry. Such symmetry is classified within the six crystal systems listed in terms of diminishing symmetry: cubic (or isometric), hexagonal, tetragonal, orthorhombic, monoclinic, and triclinic. The literature is not necessarily consistent on how many crystal systems there are, and many authors divide the hexagonal system into separate hexagonal and trigonal systems for a total of 7. This classification is used here following Auld (1990) and Tinder (2007). Regardless of the preference, nine different sets of stiffnesses $[c]$ can be defined. This is larger than the seven classes because two unique sets exist in both the trigonal and the tetragonal systems depending on the slightly different symmetries. The lower the symmetry, the greater the number of independent elastic stiffnesses required. The organization of the nine different sets of $[c]$ is shown symbolically in **Figure 3**.

Bass (1995) had compiled an extensive collection of C_{ij} for many minerals up to 1995, and Angel et al. (2009) provided an overview of the techniques used to obtain these elastic constants. The full sets of the values in $[c]$ are repeated here for a

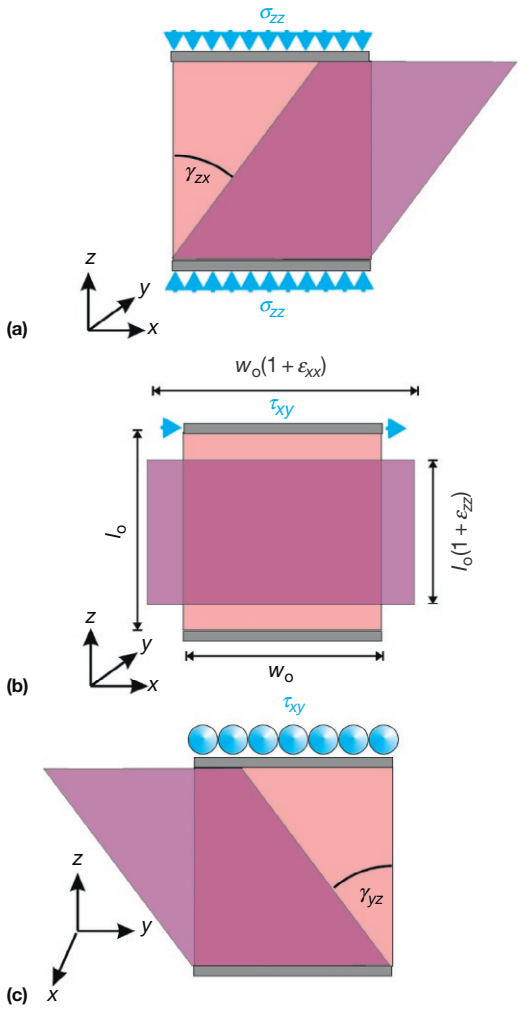


Figure 2 Illustration of the physical meaning of the elastic moduli that provide stress–strain responses outside of those defined in **Figure 1**. (a) Generation of a shear deformation of angle γ_{zx} by the application of a uniaxial stress σ_{zz} leading to the definition of $C_{53}=\sigma_{zz}/\gamma_{zx}$. (b) Generation of normal strains ϵ_{xx} and ϵ_{zz} by the application of a shear stress (shown as a pure shear) τ_{xy} leading to the definitions of $C_{16}=\tau_{xy}/\epsilon_{xx}$ and $C_{36}=\tau_{xy}/\epsilon_{zz}$, respectively. (c) Generation of a shear strain γ_{yz} by application of the shear stress τ_{xy} leading to the definitions of $C_{16}=\tau_{xy}/\gamma_{yz}$.

few of the more important minerals in **Table 1**. While most of the values are from Bass (1995), it should be no surprise that his compilation did not include any values for triclinic minerals, as determining all 21 independent elastic constants for a triclinic mineral awaited the more recent work of Brown et al. (2006).

11.03.3.2 Bounds on Isotropic Mixtures of Anisotropic Minerals

The discussion in the preceding text revealed that all minerals are elastically anisotropic to some degree. However, we often prefer that the rocks that we study, which are of course an assemblage of these minerals, be isotropic. That is, most workers prefer to indirectly employ the values in [c] distilled into bulk K and shear μ moduli for an isotropic polycrystal; it is this isotropic value that is often all that is reported in compilations. This value is usually called the Voigt–Reuss–Hill (VRH) average (Hill, 1952), being the simple arithmetic mean of the isostress Voigt (Voigt, 1928) and isostrain Reuss (1929b) bounds. The Voigt bulk K_V and shear μ_V moduli for any crystal class (Anderson, 1963; Hill, 1952) are

$$K_V=\frac{1}{9}[(c_{11}+c_{22}+c_{33})+2(c_{12}+c_{23}+c_{13})]$$
$$\mu_V=\frac{1}{15}[(c_{11}+c_{22}+c_{33})-(c_{12}+c_{23}+c_{13})+3(c_{44}+c_{55}+c_{66})] \quad [15]$$

while the corresponding Reuss bulk K_R and shear μ_R moduli are

$$K_R=\frac{1}{[(s_{11}+s_{22}+s_{33})+2(s_{12}+s_{13}+s_{23})]}$$
$$\mu_R=\frac{15}{[4(s_{11}+s_{22}+s_{33})-4(s_{12}+s_{13}+s_{23})+3(s_{44}+s_{55}+s_{66})]} \quad [16]$$

One problem can be that the mineral moduli are reported as stiffnesses in [c] but eqn [16] uses the compliances of [s]. To overcome this issue, Watt and coworkers (Watt, 1979, 1980, 1986; Watt and Peselnick, 1980) laboriously derived the expressions for K_R and μ_R in terms of the stiffnesses C_{ij}. Today, however, matrix inversions are easily accomplished and one may prefer to first calculate the compliances s_{ij} using eqn [13] and substitute them directly into eqn [16].

Berryman (2005) also provided alternative formulas developed using a self-consistent approach.

These Voigt and Reuss moduli usually provide the upper and lower bounds to the moduli, respectively, as shown by Hill (1952). The values for isotropic polycrystals that are most often reported in compilations (**Table 1**), however, are the simple VRH arithmetic mean averages of these two bounds for the bulk

$$K_{VRH}=[K_R+K_V]/2 \quad [17]$$

and the shear moduli

$$\mu_{VRH}=[\mu_R+\mu_V]/2 \quad [18]$$

The differences in the bounds are insignificant for the highly symmetrical isotropic and cubic cases. Indeed, **Table 1** purposefully displays more significant digits than legitimate just to show the minor differences between the values for pyrope garnet. However, for hexagonal ice (Ih), the differences between the bounding shear moduli differ by 40%.

More rigorous bounds may also be found. *Simmons* (Meister and Peselnick, 1966; Peselnick and Meister, 1965; Simmons and Wang, 1971; Watt, 1979, 1980, 1986) derived stronger bounds on the isotropic polycrystalline moduli using the variational principles of Hashin and Shtrikman (1962) for all of the crystal classes except for triclinic. Of course, these become increasingly complicated as symmetry decreases and more elastic constants are required; and we repeat here only Simmon's (see Simmons and Wang, 1971) expression for cubic minerals:

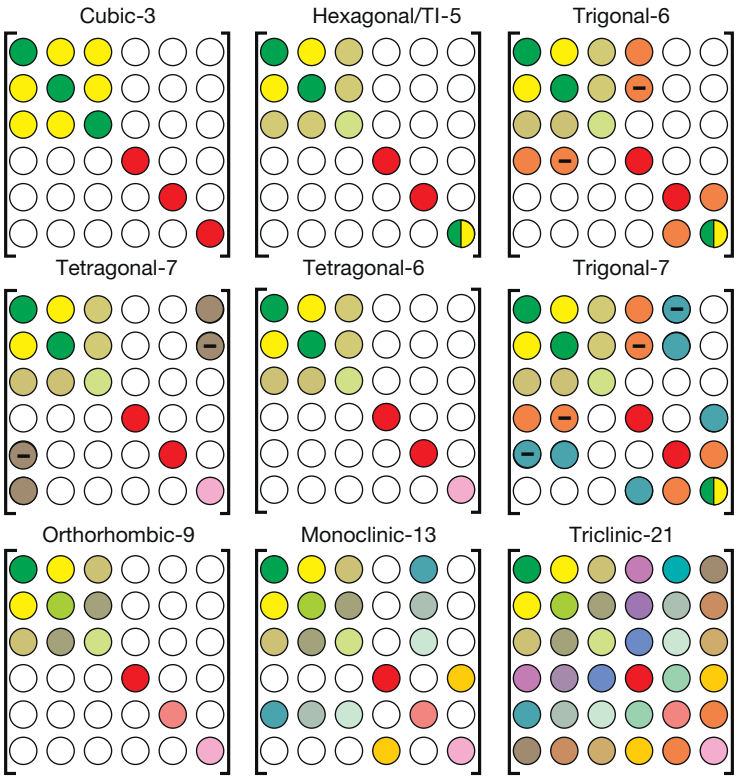


Figure 3 Graphic description of the nine different elastic stiffness matrices [c] in Voigt abbreviated notation following eqn [10] and employing Tinder's (2007) form. The symmetrical system and its corresponding number of independent C_{IJ}'s are given by the title immediately above each matrix. The values of the C_{ij} are color-coded. A white circle indicates zero. The same solid color for different components means that they share the same numerical value. A minus sign indicates that the component's value is the negative of the adjacent component of the same solid color. Note that C_{66} in the hexagonal and trigonal [c] matrices has two colors; this is to indicate that in these systems, $C_{66}=(C_{11}-C_{12})/2$.

$$
\begin{aligned}
&K_{HS^+}=K_{HS^-}=K_V=K_R=K\\
&\mu_{HS^+}=G_1+3\left[\frac{5}{G_2-G_1}+4\frac{3K+2G_1}{5G_1(3K+4G_1)}\right]^{-1}\\
&\quad\geq\mu\geq G_2+2\left[\frac{5}{G_1-G_2}+6\frac{3K+2G_2}{5G_2(3K+4G_2)}\right]^{-1}=\mu_{HS^-}\\
&\quad G_1=\frac{1}{2}(c_{11}-c_{12})\\
&\quad G_2=c_{44}
\end{aligned}
\qquad [19]
$$

Calculations of these different averages generally, but not always, show that the VRH averages lie within the HS bounds. This suggests that the VRH averages provide reasonable estimates of the isotropic polycrystal properties. The validity of the VRH averages in providing representative values still remains of current interest (Berryman, 2005, 2012; Pham, 2011).

Table 1 lists K_{VRH} and μ_{VRH} for many of the minerals commonly found within crustal rocks; these values are often used for calculating rock properties. This list is not intended to be comprehensive and many more complete compilations are available (Babuška and Cara, 1991; Bass, 1995; Gebrande, 1982; Hearmon, 1946, 1956; Huntington, 1958; Simmons and Wang, 1971; Sumino and Anderson, 1984). Many of the moduli in these compilations are reproduced from the extensive acoustic measurements made by Alexandrov and coworkers in the 1960s (Alexandrov and Ryzhova, 1961a,b, 1962, 1964; Ryzhova et al., 1966) on large (and imperfect) single crystals (see Brown et al., 2006). The impressive contributions of Alexandrov and coworkers have been invaluable to the geophysical community, but an update of their measurements on the minerals forming rocks in the upper crust using more modern techniques is long overdue.

11.03.3.3 Isothermal Versus Adiabatic Moduli

So far, the bulk modulus has been defined as the simple quotient of an applied hydrostatic stress to the resulting volumetric strain P/θ as shown in Figure 1. This is true, but one must also be careful about the conditions under which the experiment is carried out. A bulk modulus measurement can be carried out either of two ways depending on how heat energy is allowed to move during the experiment:

i. Adiabatically (also isentropically) when the material is compressed and no heat is allowed to flow in or out of the system
ii. Isothermally where heat can flow in or out in order to maintain the temperature of the system constant

Table 1 Values of the elastic stiffnesses for select materials displaying varying degrees of symmetry

'ij'	Isotropic Fused SIO2	Cubic Pyrope	hexagonal ice (Ih) (257 K)	Trigonal - 6 α-quartz	Trigonal - 7 Dolomite	Tetragonal - 6 Rutile	Tetragonal - 7 Wulfenite	Orhtorhombic Enstatite	Monoclinic Muscovite	Triclinic Low Albite
	Stiffnesses in GPa									
11	78.08	296.2	13.5	86.6	205	269	109	225	183	69.9
22	78.08	296.2	13.5	86.6	205	269	109	178	178	183.5
33	78.08	296.2	14.9	106.1	113	480	92	214	59.1	179.5
44	31.38	91.6	3.09	57.8	39.8	124	26.7	77.6	16	24.9
55	31.38	91.6	3.09	57.8	39.8	124	26.7	75.9	17.6	26.8
66	31.38	91.6	3.5	40	67	192	33.7	81.6	72.4	33.5
12	15.32	111.1	6.5	6.7	71	177	68	72.4	48.3	34
13	15.32	111.1	5.9	12.6	57.4	146	53	54.1	23.8	30.8
23	15.32	111.1	5.9	12.6	57.4	146	53	52.7	21.7	5.5
14	0	0	0	–17.8	–19.5	0	0	0	0	5.1
15	0	0	0	0	13.7	0	0	0	–2	–2.4
16	0	0	0	0	0	0	–13.6	0	0	–0.9
24	0	0	0	17.8	19.5	0	0	0	0	–3.9
25	0	0	0	0	–13.7	0	0	0	3.9	–7.7
26	0	0	0	0	0	0	13.6	0	0	–5.8
34	0	0	0	0	0	0	0	0	0	–8.7
35	0	0	0	0	0	0	0	0	1.2	7.1
36	0	0	0	0	13.7	0	0	0	0	–9.8
45	0	0	0	0	0	0	0	0	0	–2.4
46	0	0	0	0	0	0	0	0	0.5	–7.2
56	0	0	0	–17.8	–19.5	0	0	0	0	0.5
K_V	**36.24**	**172.8**	**8.722**	**38.1**	**99.4**	**217**	**73.11**	**108**	**67.5**	**63.8**
K_R			**8.717**	**37.6**	**89.1**	**209**	**71.58**	**107**	**48.7**	**55.2**
K_H			**8.72**	**37.85**	**94.25**	**213**	**72.35**	**108**	**58.1**	**59.50**
μ_V	**31.38**	**91.98**	**3.5093**	**47.6**	**51.8**	**124.6**	**19.75**	**76.2**	**43**	**41.2**
μ_R		**91.98**	**5.689**	**41**	**40**	**98.7**	**25.33**	**75.2**	**27.6**	**29.8**
μ_H		**92.0**	**4.60**	**44.3**	**45.9**	**111.7**	**22.5**	**75.7**	**35.3**	**35.5**
M_V	**78.08**	**295.4**	**13.4**	**101.6**	**168.5**	**383.1**	**99.44**	**210**	**125**	**118.7**
M_R		**295.4**	**13.3**	**92.2**	**142.4**	**340**	**105.4**	**208**	**85.4**	**94.93**
M_H		**295.4**	**13.4**	**96.9**	**155.5**	**361.6**	**102.4**	**209**	**105**	**106.8**

Cell color fills correlate to the symbolic representation of the Voigt elastic stiffness matrices in **Figure 3**. Values are taken from Bass (1995) except for isotropic fused SiO_2 (Ohno et al., 2000) and triclinic low albite (Brown et al., 2006). Values in bold are calculated from the other elastic stiffnesses.

A propagating compressional wave consists of alternating zones of high pressure and low pressure that are slightly hotter and colder, respectively, than the initial ambient temperature. An isothermal state could only be achieved if during a half-period, the heat could fully conduct from the high- to the low-pressure regions. This is in fact difficult to achieve (Condon, 1933; Fletcher, 1974, 1976) with the somewhat paradoxical result that isothermal conditions can only be reached at extremely high frequencies where, in a gas, the wavelengths are comparable to the mean free path of the molecules (Condon, 1933).

Table 2 is typical of the types of data presented in the literature for such properties but its simplicity does hide a number of complications that the reader, usually simply looking for a value to use, should still keep in mind. The first is that the values reported are usually the VRH averages as just noted.

The values reported are also usually the adiabatic bulk moduli as would be measured using acoustic techniques.

In most of the solid crystalline materials, this will insignificantly differ from the isothermal bulk modulus K_T obtained from static measurements carried out at constant temperature. Voigt (1928) (p. 788; see also Hearmon (1946)) provided the relationship between the isothermal $\overset{T}{}_{ik}$ and the adiabatic s_{ij} compliances:

$$s_{ik}^{T} = s_{ik} + \frac{\alpha_i \alpha_k T}{\rho C_P} \quad [20]$$

where ρ is the density, T is the absolute temperature in $^\circ K$, C_P is the heat capacity at constant pressure, and α_i is the appropriate component of the thermal expansion tensor (see Landau and Lifshitz, 1970, Section 10). This difference can often be ignored for most minerals, although for KCl, the two S_{12} values differ by 18%. The reader may be able to obtain appropriate values for α_i and C_P and formulas for their correction to pressure and temperature in the compilations of Holland and Powell (1998). It is important to note, however, that these differences cannot generally so easily be ignored for fluids. This issue will be discussed in more detail later.

11.03.3.4 Effects of Pressure and Temperature on Mineral Moduli

The values in Table 2 are for the most part provided under standard or room conditions. In many cases, in the upper crust, the use of such values as 'constants' is generally valid as the variations with pressure and temperature will be small. For example, the variations in the properties are usually ignored in petroleum and near-surface seismology. Those studying shallow and high-temperature geothermal systems may, however, need to take such variations into account.

Even though we can often ignore their effects, the moduli for various minerals are not constants and do depend on both the temperature and the confining pressure. With increasing confining pressure, minerals become stiffer and more rigid, and the first-order derivatives of the bulk $\delta K/\delta P$ and $\delta \mu/\delta P$ are positive and the range for most minerals is between 4–6 Pa Pa^{-1} and 0.5–2 Pa Pa^{-1}, respectively. Conversely, the moduli decrease with temperature with $\delta K/\delta T \sim -5$ MPa $^\circ K^{-1}$ to -30 MPa $^\circ K^{-1}$. As such, within the crust, temperature will usually influence the moduli more than pressure. The changes in the moduli and density of an isotropic polycrystalline quartz compact with pressure and temperature are shown for purposes of illustration in Figure 4. These suggest that the effects of pressure and temperature tend to cancel each other out, although in areas with steep thermal gradients, the changes in temperature can significantly alter the bulk modulus and the density of the quartz. The range of temperatures is intentionally shown to reflect the larger variations in the physical properties of the quartz in the vicinity of the α-quartz to β-quartz phase transition.

11.03.3.5 Mineral Densities

Equations [1] and [2] also include the material's bulk density ρ. Consider a rock composed of m separate mineral phases each with density ρ_i and ϕ_i of the fractional volume of the material (with $\Sigma\phi_i = 1$). The density for this rock takes the form

$$\rho = \sum_{i=1}^{m} \phi_i \rho_i \quad [21]$$

Our knowledge of mineral density is often taken for granted, but in reality, it may not always be so easy to determine and a number of different methods are applied. While usually the differences are small, one may still need to take care to know how the number is derived. The methods used to determine density include the following:

i. Archimedean fluid displacement in which the mineral sample's volume is found from the amount of reference fluid volume that it displaces. The density so determined is referred to as the specific gravity and it is simply equal to the ratio of the sample's measured mass to the volume of the fluid displaced.
ii. Boyle's law pyncnometry where the sample volume is determined by in a measurement of the deviation of the gas pressure in a container of a known volume (Lowell et al., 2004). The noble gas helium is usually used.
iii. x-Ray and neutron scattering crystallography that provide information on the dimensions and hence volume $V_{x\text{-ray}}$ of

Table 2 Density, isotropic bulk moduli, and isotropic polycrystalline wave speeds

Material	*Density (kg m^{-3})*	*Adiabatic bulk modulus* K *(GPa)*	*Shear modulus* μ *(GPa)*	*Poisson's ratio* ν	V_P *(m s^{-1})*	V_S *(m s^{-1})*
Tectosilicates						
α-Quartz	2648	37.8	44.3	0.079	6048	4090
β-Quartz[a]	2522	42.97	41.41	0.135	6239	4052
Albite	2610[b]	56.9	28.6	0.285	6034	3310
Anorthite	2765[b]	84.2	39.9	0.295	7049	3799
Orthoclase	2571[b]	62	29.3	0.296	6270	3376
Microcline	2567	55.4	28.1	0.283	6015	3309
Neosilicates						
Forsterite	3221	129.5	81.1	0.241	8589	5018
Fayalite	4380	134	50.7	0.332	6784	3402
$Fo_{91}Fa_9$	3325	129.5	77.6	0.250	8370	4831
Grossular	3602	168.4	108.9	0.234	9331	5498
Pyrope	3567	172.8	92	0.274	9101	5079
Inosilicates						

(*Continued*)

Table 2 (Continued)

Material	*Density (kg m^{-3})*	*Adiabatic bulk modulus* K *(GPa)*	*Shear modulus* μ *(GPa)*	*Poisson's ratio* ν	V_P *(m s^{-1})*	V_S *(m s^{-1})*
Hornblende	3150	93.3	49.3	0.275	7105	3956
Augite	3320	95	59	0.243	7233	4216
Ferrosilite	4002	101	52	0.280	6524	3605
Enstatite	3272	107.8	75.7	0.215	7987	4810
Phyllosilicates						
Muscovite (Illite?[c])	2844	58.2	35.3	0.248	6084	3523
Biotite	3050	50.5[d]	27.4[d]	0.270	5342	2997
Kaolinite	2669[b]	71.1[e]	31.2[e]	0.309	6498	3419
Na montmorillonite, dry	2700[f]	82[g]	32[g]	0.327	6795	3443
Na montmorillonite, wet	1700[f]	36[g]	16.5[g]	0.301	5841	3115
Talc[h]	2793	41.6	22.6	0.270	5068	2845
Lizardite[i]	2516	61	33.9	0.265	6400	3600
Antigorite[i]	2668	69	34.0	0.288	6520	3570
Evaporites						
Calcite	2712	73.3	32	0.309	6539	3435
Aragonite	2930	46.9	38.5	0.178	5790	3625
Dolomite	3795	94.9	45.7	0.293	6408	3470
Anhydrite	2963	54.9	29.3	0.273	5631	3145
Gypsum	2317	42	15.4	0.337	5195	2578
Halite	2163	24.9	14.7	0.253	4536	2607
Sylvite	1987	18.1	9.4	0.279	3926	2175
Barite	4473	55	22.8	0.318	4369	2258
Oxides						
Corundum	3982	253.5	163.2	0.235	10877	6402
Magnetite	5206	161	91.4	0.261	7371	4190
Periclase	3584	160	130.3	0.180	9650	6030
Ice-I (270 K)	917.5	8.73	3.4	0.328	3802	1925
Sulfides						
Pyrite	5016	142.7	125.7	0.160	7865	5006
Galena	7597	58.6	31.9	0.270	3649	2049
Sphalerite	4088	77.1	31.5	0.320	5398	2776
Others						
Diamond	3512	443	535.7	0.069	18153	12350
Graphite	2260	161.0	109.3	0.223	11650	6954
Fluids						
Air (STP) dry[k]	1.31	0.00014311	–	–	330.2	–
Water (STP)[l]	999.84	1.97	–	–	1402.4	–
Seawater (STP), 35% salinity[m]	1032.8	2.17	–	–	1449.1	–
Light oil (1 atm, 26 °C)[n]	750	1.15	–	–	1237	–
Heavy oil (1 atm, 26 °C)[n]	1037	2.71	–	–	1616	–
Tholeiitic basalt 1505 °K[o]	2650	17.9	–	–	2599	–
Andesite melt 1553 °K[o]	2440	16.1	–	–	2569	–
Rhyolite melt 1553 °K[o]	2310	13.5	–	–	2417	–

Unless otherwise indicated, the densities, adiabatic bulk moduli, and shear moduli are all from Bass (1995). The seismic velocities are calculated using eqns [1] and [2].

[a]From Table 4 of Ohno et al. (2006) at 575.5 °C.

[b]Densities from Smyth and McCormick (1995).

[c]See discussion in Cholach and Schmitt (2006).

[d]Value is the Voigt–Reuss–Hill average calculated assuming hexagonal symmetry (Watt and Peselni, 1980) using the values reported in Bass (1995).

[e]Value is the Voigt–Reuss–Hill average calculated assuming hexagonal symmetry (Watt and Peselni, 1980) using the values reported in Karmous (2011). See also Militzer et al. (2011) and Sato et al. (2005).

[f]Densities of a dry and saturated montmorillonite as updated by Chitale and Sigal (2000).

[g]Value is the Voigt–Reuss–Hill average calculated assuming hexagonal symmetry (Watt and Peselni, 1980) using the elastic stiffness calculated in Ebrahimi et al. (2012).

[h]Density and estimates of moduli in Bailey and Holloway (2000).

[i]Estimates from Auzende et al. (2006). See also Schmitt et al. (2007).

[j]From measurements described in Christensen (2004). See also recent modeling by Mookherjee and Stixrude (2009).

[k]From Lemmon et al. (2000).

[l]Online calculation Lemmon et al. (2012a) derived from the model of Wagner and Pruss (2002).

[m]From Wong and Zhu (1995) and Safarov et al. (2012).

[n]From ultrasonic measurements of Wang et al. (1990).

[o]Zero frequency (relaxed) estimates based on ultrasonic measurement of Rivers and Carmichael (1987) as reported in Bass (1995).

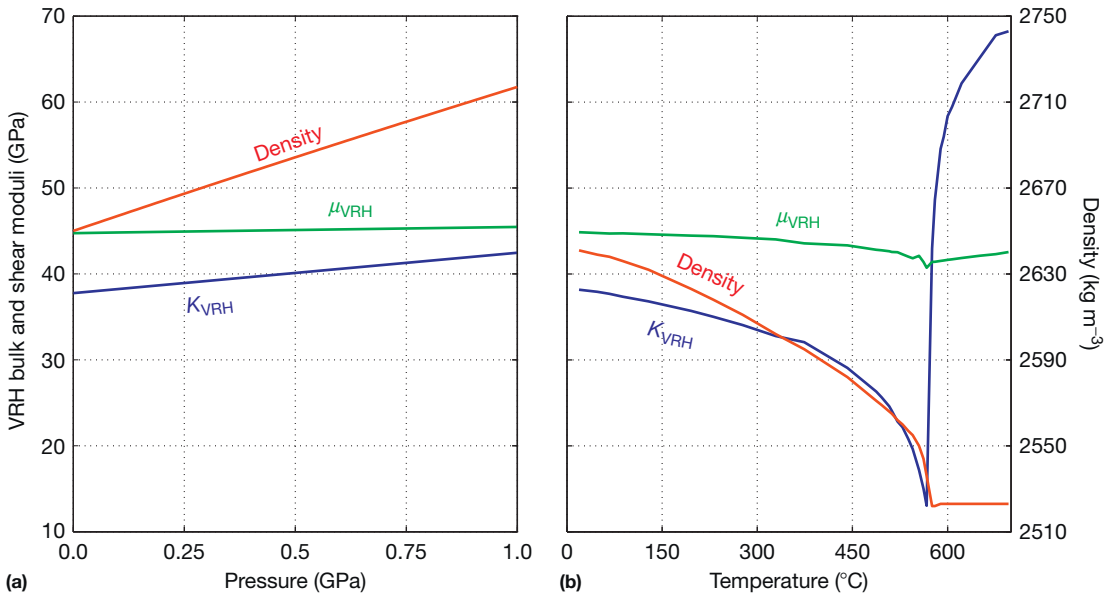


Figure 4 Changes in the averaged bulk K_{VRH} and shear μ_{VRH} moduli and density of an isotropic polycrystalline quartz aggregate (a) calculated as a function of pressure at constant temperature 298 °K (Calderon et al., 2007), The density as a function of pressure is calculated with values of $K_o = 37.7$ GPa, $dk/dP = 4.69$ Pa Pa^{-1}, and $\rho = 2650$ kg m^{-3} using the Birch–Murnaghan equation of state (Birch, 1947) (b) measured as a function of temperature at constant room pressure (Ohno et al., 2006). Note in (b) the influence of the solid state phase transition of α-quartz to β-quartz at 573.0 °C.

the mineral's crystal unit cell (Smyth and McCormick, 2013; Wohlenberg, 1982). The x-ray density $\rho_{x\text{-ray}}$ is then calculated with knowledge of the chemical molar molecular weight M in g mol^{-1} and its number of formula units per unit cell z:

$$\rho_{x\text{-ray}} = \frac{zM}{N_A V_{x\text{-ray}}} \quad [22]$$

where N_A is Avogadro's number (6.02214×10^{23}). These methods are particularly advantageous for the many minerals that do not exist in large crystal form and are then not amenable to the Archimedean or Boyle's law methods. Further, they can be used to determine density (and hence static bulk elastic properties) under conditions of pressure and temperature (Hemley et al., 2005).

iv. The electron bulk density ρ_e is calculated from estimates of the number of electrons per unit volume, n_e, from Compton scattering attenuation of x-rays (Schmitt et al., 2003):

$$\rho_e = \frac{n_p/Z}{N_A} M \quad [23]$$

where $n_e = n_p$ is the number of protons per unit volume and Z is the atomic number (i.e., number of electrons or protons in the electrically neutral molecule). This is no more than measuring the number of moles of the compound in a given volume.

The mass of a given atom is close to the sum of the masses of its constituent neutrons and protons. Neutrons and protons have nearly the same mass. Further, examination of the periodic table shows that there are equal numbers of neutron N and protons Z in the most common isotopes of elements that constitute the rock-forming minerals. It follows from this then that the ratio $M/Z \sim 2$ g mol^{-1} and eqn [23] may often be simplified to

$$\rho_e = (2\,\text{g mol}^{-1}) \frac{n_p}{N_A} \quad [24]$$

Equation [24] is often applied in geophysical well logging as it allows for an accurate estimate of the bulk density to be obtained even in the absence of detailed knowledge of the mineralogy.

Some of these measures of density are provided **Table 2**.

11.03.4 Fluid Properties

Pore fluids can strongly influence the overall seismic properties of rocks. Pore fluids contribute to the overall moduli and the bulk density of the rocks that they are contained in. Consequently, they must affect the seismic wave speeds through eqns [1] and [2]. In many respects, fluids are more interesting than solids because their density ρ_f, bulk modulus K_f, and phase state depend on pressure P and temperature T. These are the dependencies that allow, in part, for successful active seismic 'time-lapse' active seismic monitoring of fluid motions during hydrocarbon production (Bianco et al., 2010; Lumley, 2001; Schmitt, 1999) or greenhouse gas sequestration (White, 2013).

In order to forward model or to interpret observed seismic velocities through fluid-saturated rocks, one must then have a good knowledge of ρ_f and K_f with pressure P and temperature T. This P–V–T relationship, or equation of state (EOS), can take many forms depending upon the degree of sophistication required and whether it can account for phase transformations. Of course, this is a vast topic and only a cursory introduction can

be given here with particular focus on some key fluids that researchers would encounter including water, carbon dioxide, and methane. While as geophysicists we would prefer to be handed easily obtained values for the desired properties, this may not always be possible and some work may be required to obtain appropriate representative values. The reader should not necessarily look for quick answers here, but this contribution attempts to at least lead the way to the relevant literature where answers might be found. An important contribution of Batzle and Wang (1992), for example, has distilled some of these complex relations into more readily applicable formulas. Their equations have been incorporated into numerous fluid property calculators for use in seismic fluid substitution calculations; because they are so widely used, these too will be provided where appropriate. This section attempts to give some appreciation of how complicated and interesting fluid properties are, particularly relative to the more consistent minerals.

11.03.4.1 Phase Relations for Fluids

Before proceeding further, it is important to review the pressure and temperature dependencies of the fluid phase. Figure 5 shows the phase boundaries for water, carbon dioxide, and methane. Most workers are familiar with solids, liquids, and gases and the melting, sublimation, and boiling curves for the transitions between these phases. These are first-order phase transformations in that the internal energy and physical properties (Figure 6) are discontinuous. Via the latent heat of the transformation, additional energy is also released (exothermic) or absorbed (endothermic) during the phase change.

Many fluids of interest can also be in their supercritical fluid phase within the Earth. In the range of pore fluid pressures and temperatures encountered, free CO_2 and CH_4 are more likely than not to be in the supercritical regime. Supercritical H_2O exists at depth within volcanic systems, and it remains a target for geothermal power generation because of its lower viscosity but still considerable enthalpy (Fridleifsson and Elders, 2005). As illustrated in Figure 6, the transition from either gas or liquid to supercritical is continuous, and this transition is referred to as second (or higher)-order. Essentially, any surface tension between the liquid and vapor disappears under supercritical conditions, and this fluid is best characterized at the microscopic scale by rapidly fluctuating regions of density. Coherent light propagating through this fluid is scattered by the density fluctuations through a phenomenon referred to as critical opalescence. Otherwise, detecting when the fluid actually becomes supercritical can be difficult because of the lack of any discontinuity in the properties. Indeed, examination of the phase diagrams shows that it is possible to go from a gas to a liquid continuously by following a P–T trajectory around the critical point. The critical point values of temperature T_c, pressure P_c, and density ρ_c are provided for a few representative fluids in Table 3.

Note that 0.1 MPa is just <1 atmosphere of pressure and $°C = °K - 273.15°$.

11.03.4.2 Equations of State for Fluids

Again, in order to determine seismic wave speeds, eqns [1] and [2] demand knowledge of the moduli and density. Although it will become fully apparent later in Section 11.03.6, an overall saturated rock's bulk modulus K_{sat} and density ρ_{sat} will require knowledge of the fluid bulk modulus K_f and density ρ_f (or equivalently the specific volume $V_f = 1/\rho_f$). As Figure 6 shows,

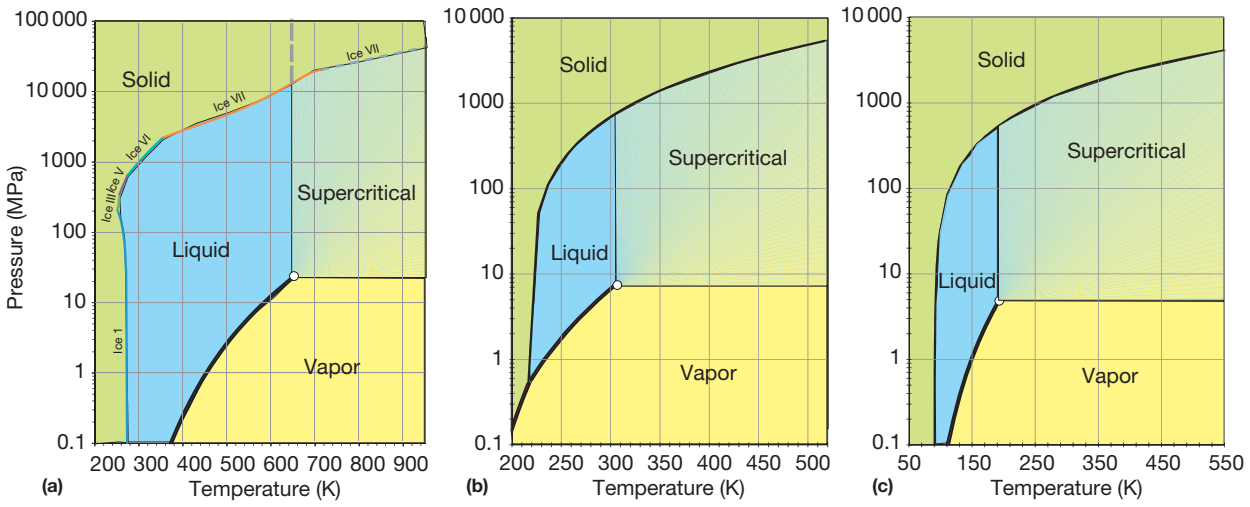


Figure 5 P–T phase diagrams for important relevant fluids. (a) Water phase diagram: The critical point shown as a white dot lies at $P_C = 22.064$ MPa and $T_C = 373.946$ °C (647.096 °K). The liquid–vapor phase boundary line is from the standard thermodynamic models (Wagner and Pruss, 2002) for water and calculated using the NIST web-based database (Lemmon et al., 2012b) and the solid–liquid/supercritical phase boundary calculated using the expressions developed by the IAPWS (2011); this boundary is only inferred past 775 K. (b) Carbon dioxide phase diagram: The critical point shown as a white dot lies at $P_C = 7.3773$ MPa and $T_C = 30.9782$ °C (304.1282 °K). The liquid–vapor phase boundary line is from the standard thermodynamic models for CO_2 (Span and Wagner, 1996) and calculated using the NIST web-based database (Lemmon et al., 2012b) and the melting and sublimation lines calculated using eqns [3.10] and [3.12] of Span and Wagner (1996). (c) Methane phase diagram: The critical point shown as a white dot lies at $P_C = 4.5992$ MPa and $T_C = -78.586$ °C (194.564 °K). The liquid–vapor phase boundary line is from eqn [3.2] and the melting curve from eqn [3.7] of Setzmann and Wagner (1991).

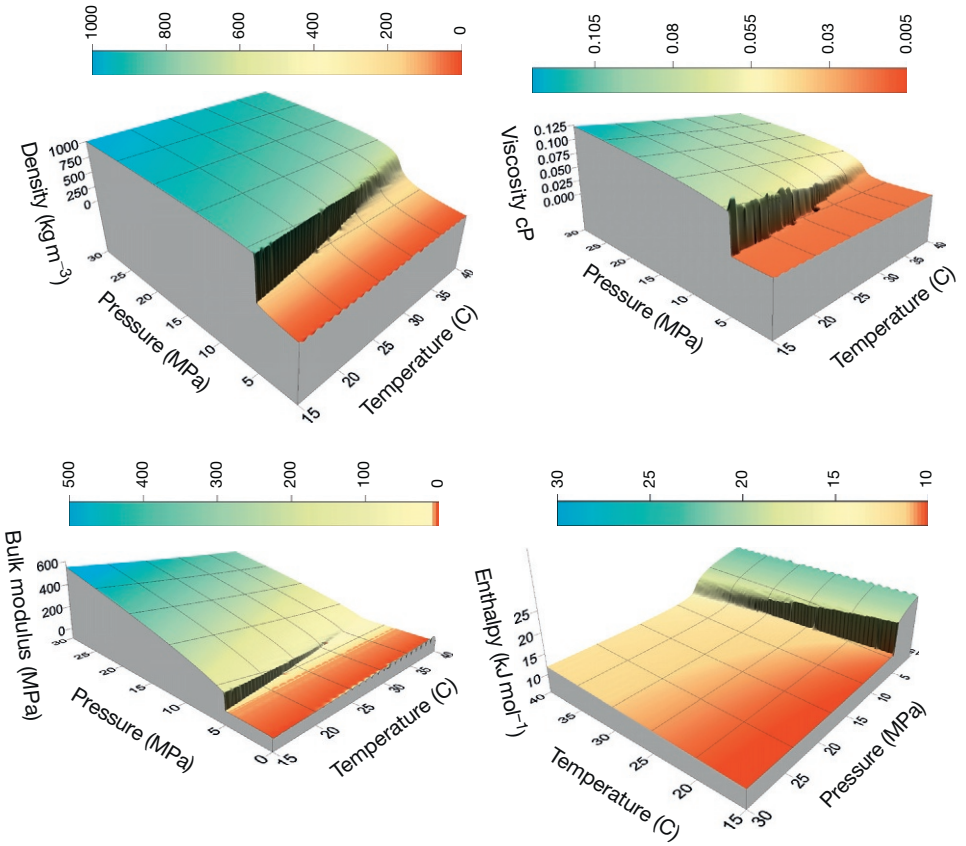


Figure 6 Dependence of the density, viscosity, bulk modulus, and enthalpy for CO_2 on pressure and temperature in the region around the critical point ($P_C = 7.3773$ MPa and $T_C = 30.9782\,^{\circ}$C (304.1282 °K)). Underlying data obtained from NIST online model (Lemmon et al., 2012b) using the model of Span and Wagner (1996).

Table 3 Thermodynamic properties of representative fluids

Molecule	*Critical temperature* T_c *(°C)*[a]	*Critical pressure* P_c *(MPa)*[a]	*Critical density (mol* m^{-3}*)*[a]	*Critical volume* ($m^3\,mol^{-1} \times 10^{-5}$)	*The van der Waals* a[b] *(Pa* ($m^3\,mol^{-1})^2$)	*The van der Waals* b[c] ($m^3\,mol^{-1} \times 10^5$)	*Acentric factor*	*Boiling point (°C)*
Water	373.946	22.064	17873.72	5.595	0.20719	1.8649	0.3443	99.974
Carbon dioxide	30.9782	7.3773	10624.9	9.412	0.19605	3.1372	0.2239	−78.4
Methane	−82.586	4.5992	10139	9.863	0.13421	3.2876	0.0114	−161.48
Propane	96.675	4.24766	5000	20.000	0.50971	6.6666	0.1524	−42.09
Decane	344.6	2.103	1640	60.976	2.34570	20.325	0.488	174.12
Benzene	288.90	4.894	3956.1	25.277	0.93810	8.4258	0.2092	80.08
Nitrogen	−146.958	3.3958	11183.9	8.941	0.08144	2.9804	0.0372	−195.79
Helium	−267.9547	0.22746	17399	5.747	0.00225	1.9158	−0.382	−268.92
Hydrogen	−240.005	1.2964	155508	0.643	0.00016	0.2143	−0.219	−252.78

[a]Data obtained from webbook.nist.gov/chemistry/fluid/.
[b]$a = 3P_cV_m^3$.
[c]$b = V_m/3$.

both are pressure- and temperature-dependent. The ρ_f maps a surface in the three-dimensional pressure–volume–temperature (P–V–T) space. This surface is mathematically described by the fluid's EOS. For our purposes, this means that ρ_f can be found if we know the *in situ* P and T. The bulk modulus K_f, however, requires the calculation of the volume-dependent derivative along the lines of either constant temperature (isotherms) or entropy (adiabats). Although in practice the simple EOS described may not always be adequate to describe real fluid behavior, it is useful to present them to provide some basic background on how density and bulk moduli may be determined. Some discussion of how one may find appropriate values is necessary, and to do this, a review of common EOS is required.

11.03.4.2.1 Ideal gas law

The simplest EOS is that for a perfect gas that considers the gas molecules to be point masses of no volume such that

$$PV_m = RT \quad [25]$$

where $R = 8.3144621(75575)\,\mathrm{J\,mol^{-1}{}^{\circ}K^{-1}}$ is the gas constant, V_m is the molar volume in $\mathrm{m^3\,mol^{-1}}$, and T is the temperature in degrees kelvin. Note that $\rho_f = 1/V_f = M/V_m$. The P–V_m relationship for a perfect gas law equation [25] is plotted for two temperatures in Figure 7, and at all temperatures, P simply decreases monotonically. Consequently, the perfect gas law is of limited value in rock physics as it only describes gas behavior at low densities and high temperatures. It cannot predict the existence of more condensed liquid or supercritical phases nor their properties.

11.03.4.2.2 Adiabatic and isothermal fluid moduli

A brief examination of eqn [25] illuminates some issues with regard to the fluid compressibility, and this brings us again to the issue of the adiabatic and isothermal bulk moduli first discussed for solids in Section 11.03.3.3. The general definition of the isothermal bulk modulus K_T for any material is

$$K_T = -V\left(\frac{\partial P}{\partial V}\right)_T = \rho\left(\frac{\partial P}{\partial \rho}\right)_T \quad [26]$$

The application of eqn [26] to the perfect gas EOS of eqn [25] yields $K_T = P$. However, as already stressed earlier, most seismic wave propagation phenomena occur adiabatically and are controlled by the adiabatic bulk modulus K_S. For an ideal gas, this is

$$K_S = -V\left(\frac{\partial P}{\partial V}\right)_S = \rho\left(\frac{\partial P}{\partial \rho}\right)_S = \frac{C_P}{C_V}K_T = \gamma K_T = \gamma P = \gamma\frac{RT}{V_M} \quad [27]$$

where the ratio of the heat capacities at constant pressure C_P and volume C_V is often called the adiabatic index γ. The value of γ depends on the degrees of freedom for the gases with $\gamma \approx 1.67$ for a monotonic gas (only three translational motions allowed) and $\gamma = 1.4$ for a diatomic gas (with two additional rotational degrees of freedom). Under adiabatic conditions, then $K_S = \gamma K_T = \gamma P$; and consequently, K_S and K_T can be considerably different for gases and many hydrocarbon fluids (Picard and Bishnoi, 1987). This translates to large differences in the wave speeds between adiabatic and isothermal conditions. Substituting $K_S = \gamma P$ into eqn [3] for the wave speed in a 'perfect' gas

$$V_L = \sqrt{\frac{\gamma P}{\rho_f}} \quad [28]$$

but more generally for nonideal fluids, one arrives at the Newton–Laplace eqn [3] rewritten here as

$$V_L = \sqrt{\frac{K_S}{\rho_f}} \quad [29]$$

For example, in dry air with $\gamma = 1.3998$ (Wong and Embleton, 1984), the adiabatic sound speed calculated using eqn [28] at 0 °C and 101.325 kPa (i.e., one standard atmosphere) is 331.29 m s^{-1} (±0.02%) in excellent agreement with

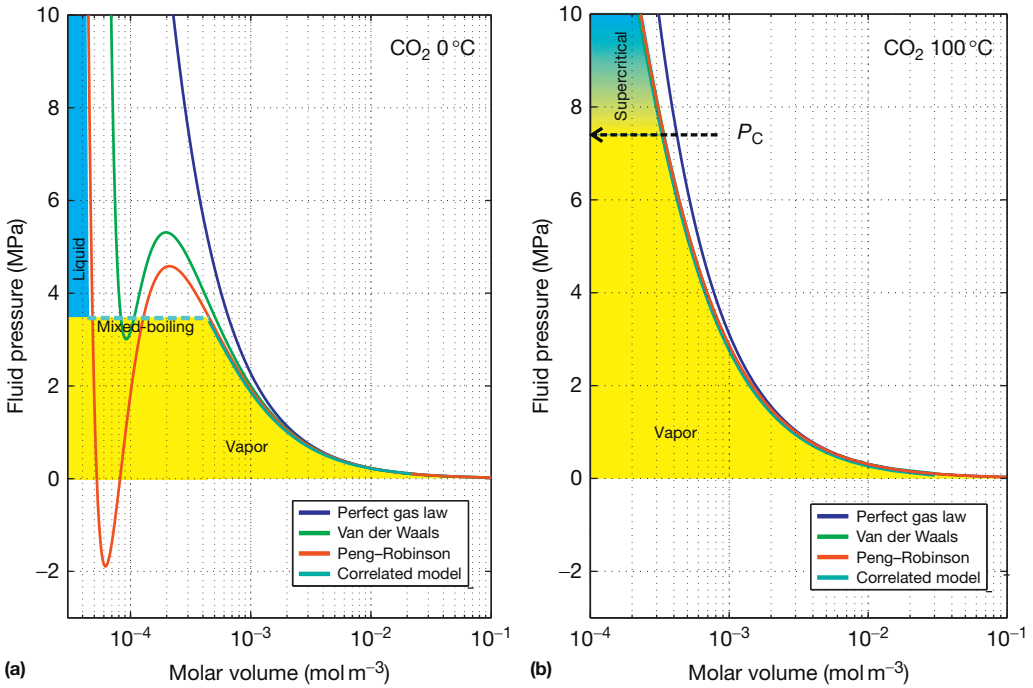


Figure 7 Comparison of different EOS for CO_2 for isotherms at (a) 0 °C and (b) 100 °C plotted in fluid pressure versus molar volume space. These include the perfect gas law eqn [25] (blue line), the van der Waals formula (eqn [32]), the Peng–Robinson equation [35] (with acentric factor = 0.22394), and the more sophisticated correlative models of Span and Wagner (1996) that directly specify phase. Filled color zones to the left of the curves indicate the various phases according to the correlative models of Span and Wagner (1996). Underlying data obtained from NIST online model (Lemmon et al., 2012b) using the model of Span and Wagner (1996).

experimental observations. In contrast, the isothermal sound speed first predicted by Newton would be only ~280 m s^{-1}.

As will be apparent later, an appropriate value for the fluid bulk modulus K_f is necessary to calculate the saturated rock properties. In the following discussions, we will assume that the fluid bulk modulus K_f that needs to be found is the adiabatic, or isentropic, bulk modulus K_S. It is worth mentioning that V_L as defined by eqn [29] is at low frequencies with the system at equilibrium a thermodynamic property in its own right; in the fluid physics and engineering communities, V_L is called the thermodynamic sound speed (Castier, 2011; Picard and Bishnoi, 1987).

Similarly to eqn [20], adiabatic and isothermal bulk moduli are related through

$$\frac{1}{K_S}=\frac{1}{K_T}-\frac{\alpha^2 T}{\rho C_P} \qquad [30]$$

where α is the coefficient of thermal expansion, T is the temperature in K, ρ is the density, and C_P is the isobaric (i.e., constant pressure) heat capacity (Clark, 1992; Kieffer, 1977).

11.03.4.2.3 The van der Waals model

The deficiencies of the perfect gas law of equation [25] were recognized early. In order to attempt to resolve these problems, van der Waals (1873) adapted eqn [25] by assigning the molecules a finite molar volume b and allowing them minor self-attractive 'van der Waals' forces to develop his EOS

$$\left(P+a/V_m^2\right)(V_m-b)=RT \qquad [31]$$

that may alternatively be displayed in its polynomial cubic form

$$V_m^3-\left(b+\frac{RT_c}{P_c}\right)V_m^2+\frac{a}{P_c}V_m-\frac{ab}{P_c}=0 \qquad [32]$$

Factor a describes the weak self-attraction of the gas molecules to one another, and a is proportional to the liquid's vaporization energy. Factor b is roughly equivalent to the molar volume of the liquid. These constants may be found from the fluid's T_c and P_c:

$$\begin{aligned} a&=\frac{(RT_c)^2}{64P_c}\\ b&=\frac{RT_c}{8P_c}\end{aligned} \qquad [33]$$

The reader will note in Figure 7 that the isotherm of the van der Waals curve at 0 °C (i.e., below the critical temperature T_c where two phases can exist at a given pressure) displays cubic behavior with three possible real roots. The smallest and largest solutions provide the molar volumes of the liquid and the vapor, respectively, while the intermediate root has no real physical meaning. At sufficiently high temperatures or at low pressures, only one real root exists. While not obvious to the more casual reader, this shape allows for the prediction of liquid, gas, and supercritical regimes. The real fluid does not follow the trajectory in the region of the function's trough and peak, which is actually a mixed-phase region where the liquid and vapor coexist until either condensation or boiling is complete. Castellan (1971) contains a particularly clear discussion of the information that may be obtained from the van der Waals curves that are beyond the scope of the discussion here. Carcione et al. (2006) had applied the van der Waals formula to obtain appropriate fluid moduli and density for use in modeling of the seismic behavior of CO_2 saturated rocks.

11.03.4.2.4 The Peng–Robinson EOS

The development of the van der Waals equation [31] was a major improvement over the perfect gas law equation [25], but as examination of Figure 7(a) shows, it is not at all accurate in the fluid state; and a plethora of additional formulas have been developed to overcome this to varying degrees of complication. The Peng and Robinson (1976) modification to the van der Waals formula

$$\left(P+\frac{\alpha_{PR}a_{PR}}{V_m^2+2b_{PR}V_m-b_{PR}^2}\right)(V_m-b_{PR})=RT \qquad [34]$$

remains one of the most popular in that it retains much of the simplicity of the van der Waals EOS but provides better estimates of the P–V–T relations. In eqn [34], there are three factors that are obtained from tabulated values that include the Peng–Robinson a_{PR} and b_{PR}

$$\begin{aligned} a_{PR}&=0.45724\frac{(RT_c)^2}{P_c}\\ b_{PR}&=0.07780\frac{RT_c}{8P_c}\end{aligned} \qquad [35]$$

and the acentric factor ω that accounts for nonsphericity of the molecule (Bett et al., 2003). The effects of the acentricity and temperature are included in α_{PR}:

$$\alpha_{PR}=\left[1+\left(0.37464+1.5422\omega-0.26992\omega^2\right)\left(1-\sqrt{T/T_c}\right)\right]^2 \qquad [36]$$

The cubic form of the Peng–Robinson equation [34] is most often for simplicity given in terms of the compressibility factor $Z=PV_m/RT$ with

$$Z^3-(1-B)Z^2+\left(A-3B^2-2B\right)Z-\left(AB-B^2-B^3\right)=0 \qquad [37]$$

where

$$\begin{aligned} A&=\frac{a_{PR}P}{(RT)^2}\\ B&=\frac{b_{PR}P}{RT}\end{aligned} \qquad [38]$$

and, as for the van der Waals equation, in the two-phase region below T_c, the largest and smallest roots correspond to the vapor and liquid molar volumes, respectively.

The Peng–Robinson equation [34] predicts well the P–V–T relationship for in the vapor and supercritical regimes. The liquid properties are not as precise but depending on the application may be sufficient. In Figure 7, for example, the Peng–Robinson EOS value for V_m at the boiling point exceeds that observed by about 7%. It too shows cubic behavior below T_c and the same arguments used for the van der Waals equation [32] hold. The Peng–Robinson equation [34] was a seminal development, and a very large literature of adaptations and additional corrections to it has emerged. It is particularly heavily used in the petroleum industry to describe P–V–T relations of both *in situ* and produced hydrocarbons.

11.03.4.2.5 Correlative EOS models

The best models, at least for monomolecular fluids, are those constructed by fitting of actual observed experimental measurements. As such, these correlative models rely on the existence of solid experimental observations of P–V–T relations and phase boundaries and of course can only be as good as the quality of the data and are limited by the ranges of conditions over which measurements have been made (Setzmann and Wagner, 1989). Models exist for important geophysical fluids including water (Wagner and Pruss, 2002), CO_2 (Span and Wagner, 1996; see also Han et al. (2010, 2011)), and methane (Kunz and Wagner, 2012; Setzmann and Wagner, 1991). The formulas used in these models are extensive and are not repeated here as they are lengthy and dependent on the phase state. The correlative model for CO_2 is shown as the 'reference' against which the other simpler models are compared in **Figure 7**. Online calculation of isotherms (P–V–T paths of constant temperature), isobars (P–V–T paths of constant pressure), and isochors (P–V–T paths of constant volume or density) is available from Lemmon et al. (2012b) for a number of important fluids. Independent software for water properties may also be found for water (NIST, 2010b) and for other fluids and mixtures of interest to geophysics (NIST, 2010a).

Pure water has been extensively studied and numerous less complicated expressions exist with validity over more limited ranges of P and T and a large literature exists. A few of the contributions in this area for liquid water in include those of Del Grosso and Mader (1972) and Kell (1975).

Batzle and Wang (1992) provided a number of simplified equations for the acoustic properties of water based on earlier experimental work of many authors (Helgeson and Kirkham, 1974; Rowe and Chou, 1970; Wilson, 1959). For example, based on Rowe and Chou's (1970) compilation, they derived a formula for the density of pure water ρ_w (in g cm^{-3}) as a function of pressure P (in MPa) and temperature T (in °C)

$$\rho_w = 1 + 10^{-6}(-80T - 3.3T^2 + 0.00175T^3 + 489P - 2TP + 0.016T^2P - 1.3\times10^{-5}T^3P - 0.333P^2 - 0.002TP^2) \quad [39]$$

and provided Wilson's (1959) empirical regression for the speed of sound V_w (longitudinal wave in m s^{-1}) in pure water

$$V_w = \begin{bmatrix}1 & P & P^2 & P^3\end{bmatrix}\begin{bmatrix}w_{00} & w_{10} & w_{20} & w_{30} & w_{40}\\ w_{01} & w_{11} & w_{21} & w_{31} & w_{41}\\ w_{02} & w_{12} & w_{22} & w_{32} & w_{42}\\ w_{03} & w_{13} & w_{23} & w_{33} & w_{43}\end{bmatrix}\begin{bmatrix}1\\ T\\ T^2\\ T^3\\ T^4\end{bmatrix} \quad [40]$$

with the coefficients w_{ij}

$$\begin{bmatrix}w_{00} & w_{10} & w_{20} & w_{30} & w_{40}\\ w_{01} & w_{11} & w_{21} & w_{31} & w_{41}\\ w_{02} & w_{12} & w_{22} & w_{32} & w_{42}\\ w_{03} & w_{13} & w_{23} & w_{33} & w_{43}\end{bmatrix} = \begin{bmatrix}1402.85 & 4.871 & -0.04783 & 1.487\times10^{-4} & -2.197\times10^{-7}\\ 1.524 & -0.0111 & 2.747\times10^{-4} & -6.503\times10^{-7} & 7.987\times10^{-10}\\ 3.437\times10^{-3} & 1.739\times10^{-4} & -2.135\times10^{-6} & -1.455\times10^{-8} & 5.230\times10^{-11}\\ -1.197\times10^{-5} & -1.628\times10^{-6} & 1.237\times10^{-8} & 1.327\times10^{-10} & -4.614\times10^{-13}\end{bmatrix} \quad [41]$$

Wilson's equation [40] may be adequate for fluid substitution use under field conditions (Chen and Millero, 1976), whereas in the laboratory, the more recent correlation models (Lemmon et al., 2012a; Wagner and Pruss, 2002) or more recent formulas (Belogol'skii et al., 1999; Lin and Trusler, 2012; Vance and Brown, 2010) may be preferable if greater accuracy is desired. That said, the adiabatic bulk modulus predicted using eqns [39] and [40] from Batzle and Wang (1992) agrees well with other models (**Figure 8**).

11.03.4.2.6 Determining K_f from equations of state

Ideally, the best way to obtain the adiabatic fluid bulk modulus K_f is from direct determinations of the sound speed and density in the fluid subject to the appropriate P–T conditions (Clark, 1992; Picard and Bishnoi, 1987). Equations [26] and [27] provide the definitions of the isothermal K_T and adiabatic K_S bulk moduli as the partial derivative of P with V_m. In principle then, one may simply obtain K_T or K_S by appropriately differentiating, respectively, either an isotherm (such as shown in **Figure 7**) or a corresponding adiabat. This is done relatively easily for the isotherms for which we can write explicit equations as in the preceding text, and, ignoring imperfections in the EOS itself, K_T can then be easily obtained. Unfortunately, for purposes of seismic wave propagation, K_S is the one required as discussed in **Section 11.03.3.3**. From the EOS, K_S can be determined in different ways:

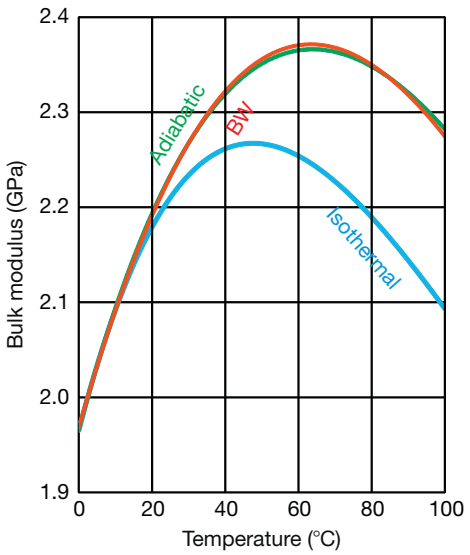


Figure 8 Comparison of the bulk moduli of liquid water at 1 atmosphere from 0 to 100 °C for (i) isothermal K_T (blue line) calculated using expressions in Kell (1975), the adiabatic K_S (green line) as calculated from the correlative model sound speeds and densities from Lemmon et al. (2012b), and a second adiabatic K_S (red line) calculated from the Batzle and Wang (1992) formulas for density (eqn [40]) and sound speed (eqn [41]). Note that the peak in the bulk moduli arises from the unique behavior of water.

i. Assuming that the differences are small such that $K_S \approx K_T$. This may be true in some cases, but even for liquid water at 1 atmosphere, the two deviate above 20 °C (**Figure 8**). This assumption becomes even more problematic for hydrocarbon fluids. In practical interpretation of seismic data, the errors in determining seismic velocities may be greater than the differences produced by using K_T, but workers should still take care when assuming K_T is sufficient.
ii. Obtaining K_T by derivation of the isotherm and then applying eqn [30] assuming appropriate knowledge of C_P and α. One problem with this approach is that the knowledge of these parameters may be incomplete or erroneous.
iii. Numerically converting the P–V isotherms to P–V isentropes and taking the derivative of these directly (Picard and Bishnoi, 1987). Again, this approach can be fraught with additional error. Clark (1992) noted that this may be the only way to obtain a value for K_S particularly for complex multiphase fluids but it should be recognized that the EOSs are usually imperfect (as demonstrated in **Figure 7**). She recommended that, if possible, workers should find K_S by measuring the sound speeds in the fluids directly with eqn [29]. Plantier et al. (2008) also echoed this caveat on the direct application of thermodynamic relations to predict wave speeds or bulk moduli in heavy oils.

An alternative strategy to calculate the isentropic compressibility and hence the fluid's thermodynamic sound speed has recently been developed by Nichita et al. (2010). They obtained the relevant parameters from partial derivatives of the total enthalpy with respect to temperature at constant pressure and composition.

11.03.4.3 Mixtures and Solutions

The discussion in the preceding text focuses primarily on simple fluids with a single chemical component, but in the Earth, this will rarely be the case. Pore waters are never pure and usually contain numerous solvated ions and absorbed gases (van Weert and van der Gun, 2012). Hydrocarbon-saturated rocks actually contain a complex mixture of different organic species mixed both miscibly and immiscibly, with water and free gas often residing separately from the oil in the same pore. Perhaps most importantly, the pore space can hold the fluid in pockets of liquids and gas, a mixture with unexpected mechanical properties (the so-called 'patchy' saturation). Consequently, we must look at the topic of fluid mixtures from a number of different perspectives as solutions, mixtures of immiscible fluids, and mixed phases.

11.03.4.3.1 Frozen mixtures

'Frozen' fluid mixtures can consist of either (i) a mix of immiscible and nonchemically interacting fluids or (ii) a mix of vapor and liquid phases of the same fluid, the proportions of which do not vary during passage of the seismic wave. In his classic text, Wood (1955) developed the formula for the wave speed in a bubbly fluid with liquid and gas phases in proportions of ϕ_l and ϕ_g, respectively, and of bulk moduli K_l and K_g, respectively (Wilson, 2005),

$$V_f = \sqrt{\frac{1}{\left(\phi_l \rho_l + \phi_g \rho_g\right)\left(\phi_l / K_l + \phi_g / K_g\right)}} \qquad [42]$$

where he assumed no chemical interaction between the phases and that the Laplace surface tension pressure in the gas bubbles was negligible (i.e., the bubbles were sufficiently large approximately >0.1 mm in water). By analogy to eqn [29], the bulk modulus is

$$K_f = \left(\phi_l / K_l + \phi_g / K_g\right)^{-1} \qquad [43]$$

In the rock physics literature, eqn [43] has been extended to mixtures of n fluids and further generalized to include suspended solids within the liquid. It is often assumed to be an exact representation:

$$K_f = \left(\sum_{i=1}^{n} \phi_i / K_i\right)^{-1} \qquad [44]$$

For purposes of illustration and following Wood's example, consider a mixture of air and water. The density of the mixture increases linearly from that for air to that for water but its bulk modulus remains much closer to that for air and only rises to that for water in at near 100% water content (**Figure 9(a)**). Conversely, a few bubbles within the water have a large influence on diminishing the mixture's bulk modulus. This leads to the unexpected result that the wave speed in the mixture is substantially less than that in either air or water (**Figure 9(b)**), a consequence of the competition between stiffness and density in eqn [3].

As noted, the results of **Figure 9** obtained using Wood's equation [42] are presumed to be valid at low frequencies, but the formula is applied nearly universally in the geophysical literature. Nichita et al. (2010) had recently questioned even this assumption. In reality, the physics of the problem is quite complex with numerous additional considerations needing to be accounted for including changes to the bulk modulus of the gas due to Laplace internal pressure in the bubbles (e.g., Kieffer, 1977; Landau and Lifshitz, 1959) and of dispersion due to a wide variety of thermal and mechanical factors (Carey and Evans, 2011), the literature of which is too large to review here.

One of the more popular models that considers wave speed dispersion is that of Commander and Prosperetti (1989). This model includes the bubble size and corresponding resonance frequency, thermal diffusivity, and the Laplace bubble pressure. Because of these factors, this model is able to determine both mixture wave speed and attenuation with frequency. This model, too, has another unexpected result that, while Wood's equation [42] appears to predict the velocities well for this case below a few hundred Hz, above the resonance frequency for the bubbles, the stiffness of the mixture increases dramatically with surprisingly unexpected wave speeds above 15 000 m s^{-1} (**Figure 10**). Cheyne et al. (1995) provided a heuristic description of this phenomenon. The minimum near 3 kHz is at the bubble resonance frequency; and the subresonant ('Wood's regime') and superresonant regimes will shift left and right as the mean bubble dimensions increase and decrease, respectively. The point of showing these phenomena is not so

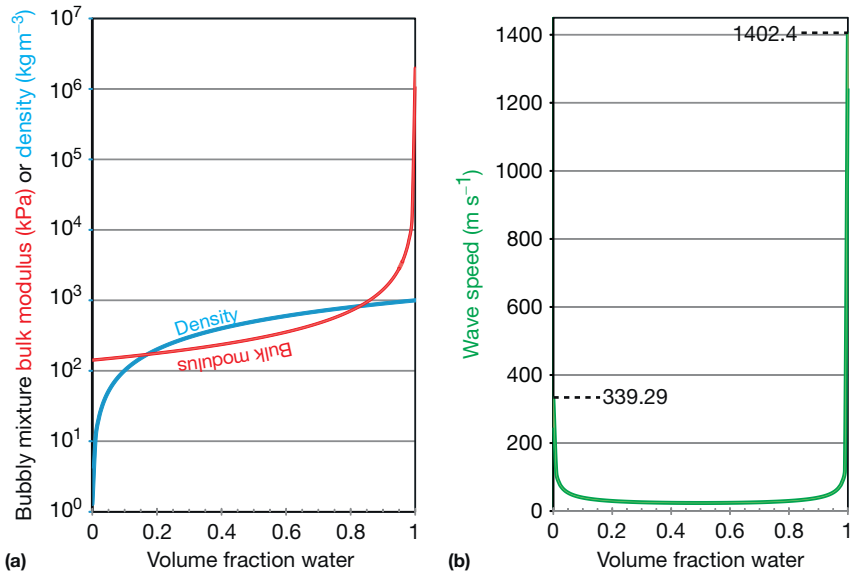


Figure 9 Properties of a bubbly mixture of air and water as a function of the volume fraction of water at 101.325 kPa and 0 °C. (a) Mixture density (eqn [21]) and bulk modulus (eqn [44]) on logarithmic scale. (b) Corresponding longitudinal wave speed in the bubbly mixture with the limits for air (339.29 m s^{-1}) and water (1402.4 m s^{-1}) shown. Calculations assume $K_{air} = 141.8$ kPa, $K_w = 1.966$ GPa, $\rho_{air} = 1.232$ kg m^{-3}, and $\rho_w = 999.84$ kg m^{-3}.

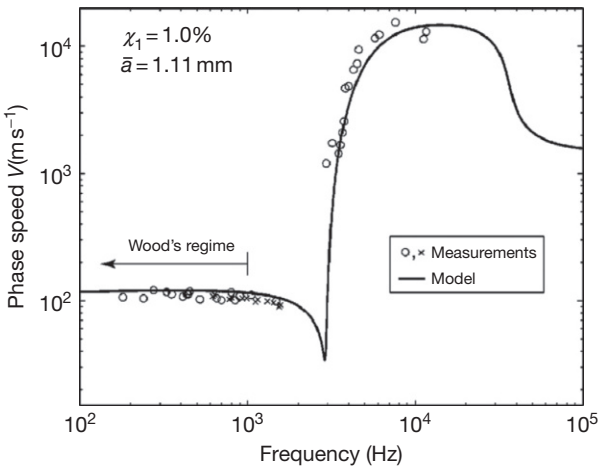


Figure 10 Observed (Cheyne et al., 1995; Wilson, 2005) and calculated (Commander and Prosperetti, 1989) wave speed dispersion in a mixture for a mean bubble radius of 1.11 m and a gas fraction air = 0.01. Reproduced from Wilson PS (2005) Low-frequency dispersion in bubbly liquids. *Acoustics Research Letters Online (ARLO)* 6: 188–194.

much that we may be interested in the wave speed of the mixture itself but that the fluid mixture's bulk moduli can be frequency-dependent. Consequently, some care may need to be exercised in the use of the otherwise ubiquitously applied general Wood's equation [44]. It is also important to note that dispersion exists even in the low-frequency Wood's regime but this is masked by the logarithmic scaling in **Figure 10**.

11.03.4.3.2 Miscible fluid mixtures

Miscible fluid mixtures are also very common within the Earth. Natural oils, for example, are really a complex blend of a multitude of different hydrocarbon compounds, and the 'heavier' (denser) and more viscous the oil, the greater the number of compounds that can occur in it. Brines are by far the most common liquids in both sedimentary basins and the crystalline crust, and their properties differ from those for pure H_2O. Molecules that would by themselves be gases, particularly CO_2 or CH_4, go into solution in both H_2O and oils. In this section, we will briefly review a number of these different types of miscible mixtures.

11.03.4.3.2.1 Brines

Solutions of water and, mostly, NaCl are the predominant fluids within the Earth's crust in both sedimentary basins and crystalline hard rocks. Brines have engendered a large literature primarily because of the need for information by a number of disciplines on their physical properties. Work on this topic continues to the present day with a standardized formula provided (IAPWS, 2008) and an alternative scheme developed by Del Grosso (1974) being the most popular descriptions of brine behavior. While the salt composition of seawater may remain globally consistent, brines in the Earth can contain numerous different electrolytes with wide concentrations affecting both the liquid density and bulk modulus. The density of the solution invariably increases with salinity, often given as some fractional measure such as ppt (referring to g of salt per kg of pure water) as illustrated for a 30 ppt brine compared with pure water in **Figure 11**. The trough in the bulk modulus near 50 °C is again apparent.

Numerous simple polynomial expressions describe the sound speeds and densities of seawater as a function of depth, salinity, and temperature (Coppens, 1981; Mackenzie, 1981; Safarov et al., 2009). Following from eqns [39] and [40], Batzle and Wang (1992) constructed fits of existing data for brine density ρ_B

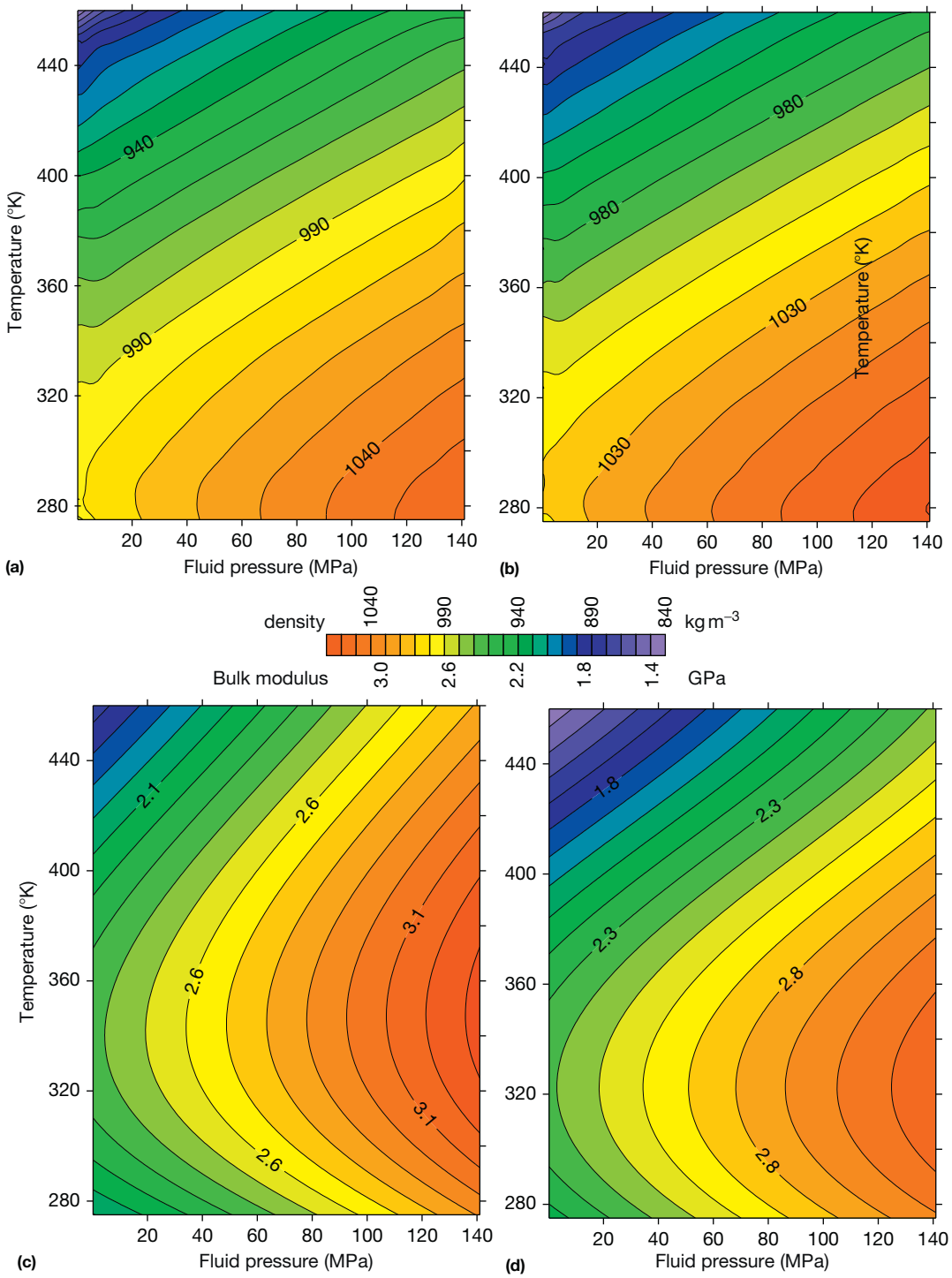


Figure 11 Comparison of the density of (a) pure water (Wagner and Pruss, 2002) and of (b) brine with a salinity of 30 ppt (g salt per kg water) (Safarov et al., 2013) contoured over the (P, T) region of 275–460°K and 0.2–140 MPa. Comparison of the (c) adiabatic bulk modulus of pure water with the (d) isothermal bulk modulus of the same 30 ppt brine (Safarov et al., 2013) over the same pressure and temperature conditions. Note that based on knowledge of the specific heat capacity of seawater (Sharqawy et al., 2010), the differences in the adiabatic and isothermal bulk moduli for seawater are significant for brines also.

$$\rho_B = \rho_w + S1\left(0.668 + 0.44S + 10^{-6} \times [300P - 2400PS + T(80 + 3T - 3300S - 13P + 47PS)]\right) \quad [45]$$

and sound speed V_B

$$V_B = V_w + S\left(1170 - 9.6T + 0.055T^2 - 8.5 \times 10^{-5}T^3 + 2.6P - 0.0029TP - 0.0476P^2\right) + S^{1.5}\left(780 - 10P + 0.16P^2\right) - 820S^2 \quad [46]$$

where S is the salinity in terms of the weight fraction in ppm/10^6. They noted that these were for NaCl concentrations and could be in significant error for other electrolytes particularly with divalent anions.

The lack of information on the properties of more concentrated brines is an important gap in our knowledge. The extensive work of Safarov et al. (2013) only reaches 30 ppt, but *in situ* natural brines can reach full salt saturations with salinities approaching 300 ppt. Data at higher concentrations are more difficult to find, but some appropriate density information might be obtained to 5.5 molal (320 ppt) in Rogers and Pitzer (1982), Gucker et al. (1975), and LaLiberte and Cooper (2004). Formation waters in the Earth too also can have quite different compositions than seawater, with a wide variety of ions including the cations Na^+, Ca^{2+}, Mg^{2+}, Fe^{2+}, K^+, Ba^{2+}, Li^+, and Sr^{2+} and anions Cl^-, SO_4^{2-}, HCO_3, CO_3^{2-}, NO_3^-, Br^-, I^-, and S^{2-}. Kumar (2003) measured wave speeds in KCl solutions but there is little in the literature on the effect of these ions on the acoustic properties of the solution. These few citations are in no way intended to be exhaustive, and there are many contributions that discuss the properties of brines. However, most of the work is performed under conditions, particularly of pressure, that are below what is needed *in situ*, and workers may need to take care in regard to the range of applicability of a given formulation.

11.03.4.3.2.2 Multicomponent fluid mixtures

Multicomponent mixtures in this section refer to miscible combinations of different gases and liquids, that is, a single-phase solution that is formed from two or more fluids. Carbonated water is one example. Natural hydrocarbon oils are another; they are quite complex and will literally contain thousands of different molecules. Indeed, we expect that such mixtures are likely the rule, not the exception, within the Earth. Such mixtures have many of the same characteristics as the pure fluids described in the preceding text, but with significant complications.

In addition to pressure P, volume V, and temperature T, the proportions x_i of the different fluids in the mixture and their respective solubilities must also now be considered. As such, the literature on the study of the behavior of such mixtures is immense and the field remains an active area of research.

Consider a mixture of two fluids A and B. The phase diagram of this multicomponent system in Figure 12 is broadly similar to that for the single components shown in Figure 5 except for the zone in P–T space separating the pure fluid and the vapor phases from one another. For the pure compounds of Figure 5, this boundary is only a line. For a miscible mixture in Figure 12, the boundary instead becomes an area in P–T space in which both liquid and vapor coexist. This

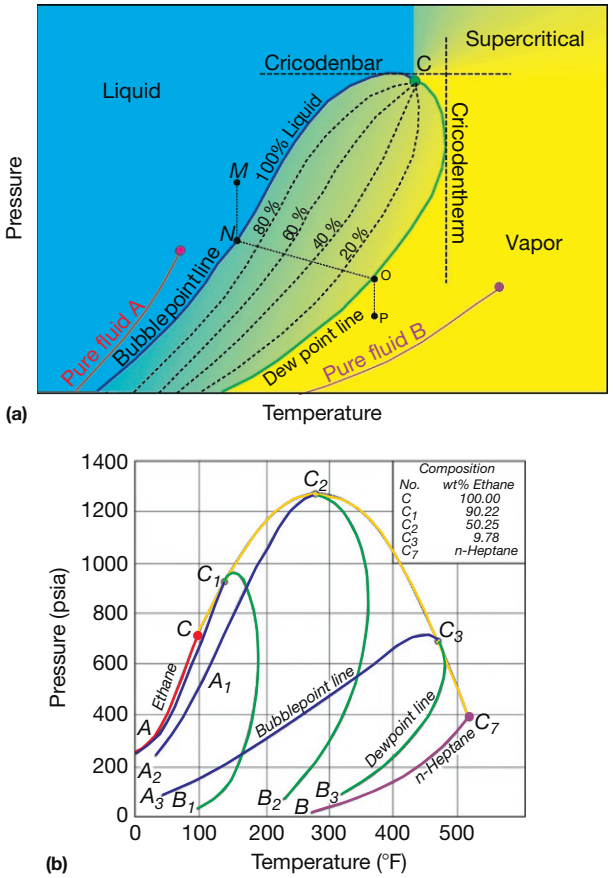


Figure 12 (a) Illustration of the phase boundaries for a 50–50 multicomponent mixture of pure fluids *A* and *B* with liquid, vapor, and supercritical regions adapted from Ezekwe (2010). The boiling lines and critical points for the two pure fluids are shown in red and green for fluids *A* and *B*, respectively. Point C is the pseudocritical point for the mixture. Percentages represent the saturation proportions of liquid to gas in the mixed-phase region between the bubble and dew point lines. (b) Mixed-phase envelopes for various proportions of ethane and *n*-heptane. The single-component boiling lines for ethane and *n*-heptane are shown in red and purple, respectively, in analogy to (a). The phase curves for three differing mixture compositions 1, 2, and 3 are shown with the bubble and dew point lines in blue and green and with their corresponding pseudocritical point. The yellow line is an envelope of the critical points and highlights the compositional dependence of the mixture's critical points. Adapted from Ezekwe N (2010) *Petroleum Reservoir Engineering Practice*. New York: Prentice Hall; Figure 4.5.

mixed-phase zone lies between the single-component boiling lines of the two pure fluids *A* and *B*.

For purposes of illustration, it is useful to follow the changes of a 50–50 mixture of two fluids along a hypothetical path *M*–*N*–*O*–*P* in *P*–*T* space in Figure 12. At *M*, the mixture is a homogeneous liquid. It remains a homogeneous liquid with decreasing pressure until point *N* is reached. Here, the path meets the bubble point line whereupon the first trace of vapor phase appears. From the appearance of this first gas bubble at *N*, the phase transformation progresses with varying saturations (i.e., proportions of liquid to gas) as one travels from *N* to *O*. At *O*, called the dew point, the last vestige of liquid phase remains. Past the dew point *O* and downward in pressure to *P*, the fluid is a homogenous vapor.

The pseudocritical point C in Figure 12 (Kay, 1936) lies where the bubble and dew point lines meet at pressure P_{PC} and temperature T_{PC}. Note that C lies above the critical points for pure A and B. The envelope of the mixed-phase boundary (i.e., the line that includes both the bubble and dew point lines) further complicates the phase behavior. Unlike the single-component fluid, C is not the limiting point at which multiple phases can exist. Consequently, the cricondenbar and the cricondentherm are the pressure and temperature limits, respectively, at which both liquid and vapor coexist. At pressure above the cricondenbar, the mixture can only be in liquid or supercritical states. At temperatures above the cricondentherm, it must be vapor or supercritical.

The location of C and the shape of the mixed-phase envelope change with the miscible fluid composition as illustrated for a real mixture in Figure 12(b). The behavior must approach that of the pure end-member as the proportions of the fluids move towards 100%. The P–T phase diagrams for various proportions of the hydrocarbon mixture of ethane (C_2H_6) and n-heptane (C_7H_{16}) show the evolution of the envelope as the proportions of the two fluids vary. The position of the pseudocritical point relative to the cricondenbar, too, varies. The yellow line in Figure 12(b) maps the locus of the pseudocritical points. It is worth noting that for the ethane/n-heptane mixture, this line is continuous between the critical points for pure ethane and n-heptane. Vankonynenburg and Scott (1980) classified this as a class I binary phase diagram (Gray et al., 2011).

The class I phase diagram of Figure 12 exemplifies the behavior of mixtures of organic liquids and particularly natural hydrocarbon oils. As noted, this is a vast topic in the hydrocarbon energy industry as such knowledge is required at nearly every point from initial production to refining. An understanding of such curves is necessary for purposes of seismic monitoring over oil reservoirs. For example, the fluid pressure in an oil field declines as fluids are produced. If this pressure decline followed a hypothetical path downward leading from point M, the pressure must eventually reach the bubble point where vapor comes out of solution. The fluid is now mixed-phase, and its overall bulk modulus drastically decreases similar to the behavior for air bubble water as shown in Figure 9. As will be seen later, this change in the fluid properties affects the overall saturated bulk modulus with a consequent decrease in both the rock's wave speed and elastic impedance. This changes the overall reflectivity of the structure with a detectable change (Fereidoon et al., 2010). Similarly, heating of liquid heavy oils to lower their viscosity and enhance their producibility would also result in the bubble point of the hydrocarbon mixture being reached with the consequential change in the seismic responses.

There are a number of resources that practitioners can exploit to obtain appropriate moduli and density. The GERG-2008 (Kunz and Wagner, 2012) model calculates the properties for hydrocarbon mixtures and it includes 21 different common components of natural gas. The software that carries out these calculations is available (Wagner, 2013). Although this is a sophisticated model, its authors still argue that new experimental data are necessary.

Batzle and Wang (1992) also provided useful approximate formulas for hydrocarbon mixtures. In their developments, a gas mixture is characterized simply only by its specific gravity G, the ratio between the gas mixture density ρ_g and that of air ρ_{air}

$$G = \frac{\rho_g}{\rho_{air}} \quad [47]$$

measured at 1 atmosphere of pressure and 15.6 °C. Following Thomas et al. (1970) and using a Benedict–Webb–Rubin EOS model (Benedict et al., 1942, 1951), they defined pseudo-reduced temperature T_{PR} and pressure P_{PR}, both of which can be related to G according to

$$P_{PR} = P/P_{PC} = P/(4.892 - 0.4048G) \quad [48]$$

$$T_{PR} = T^K/T^K_{PC} = T^K/(94.72 + 170.75G) \quad [49]$$

where the superscript K in eqn [49] indicates that the temperatures must be the absolute temperatures in degrees kelvin (T^K (°K) = T (°C) + 273.15). With these variables, the gas mixture's density ρ_g at P and T is approximately given by

$$\rho_g \cong \frac{28.8GP}{zRT^K} \quad [50]$$

where z is

$$z = \left[0.03 + 0.00527(3.5 - T_{PR})^3\right]P_{PR} + 0.642T_{PR} - 0.007T^4_{PR} - 0.52 + 0.109(3.85 - T_{PR})^2 \exp\left\{-\left[0.45 + 8(0.56 - 1/T_{PR})^2\right]P^{1.2}_{PR}/T_{PR}\right\} \quad [51]$$

The adiabatic bulk modulus K_S for this mixture may also be estimated as

$$K_S \cong \frac{P}{\left(1 - \frac{P_{PR}}{Z}\frac{\partial z}{\partial P_{PR}}\right)_T} \times \left\{0.85 + \frac{5.6}{P_{PR} + 2} + \frac{27.1}{(P_{PR} + 3.5)^2} - 8.7\exp[-0.65(P_{PR} + 1)]\right\} \quad [52]$$

with the derivative $\partial z/\partial P_{PR}$ taken at constant temperature calculated from eqn [51].

Batzle and Wang (1992) also gave expressions for estimating the density and bulk modulus of oils. As already noted, natural petroleum oils are composed of numerous different organic compounds, and, as was done with natural gas in eqn [47], these natural oils can be characterized by their density ρ_o, again measured at 15.6 °C and atmospheric pressure. They developed expressions for the pressure and the temperature dependence of the density of such oils (in g cm^{-3})

$$\rho = \frac{\rho_o + (0.00277P - 1.71 \times 10^{-7}P^3)(\rho_o - 1.15)^2 + 3.49 \times 10^{-4}P}{0.972 + 3.81 \times 10^{-4}(T + 17.78)^{1.175}} \quad [53]$$

that lead to the adiabatic bulk modulus

$$K_S \cong \rho\left\{2096\left(\frac{\rho_o}{2.6 - \rho_o}\right)^{1/2} - 3.7T + 4.64P + 0.0115 \times \left[4.12\left(\frac{1.08}{\rho_o} - 1\right)^{1/2} - 1\right]TP\right\}^2 \quad [54]$$

Live oils are oils containing gas in solution. As Figure 12 suggests, this gas will come out of solution upon depressurization at the bubble point N; and consequently, the composition and physical properties of a produced oil may be significantly different from its *in situ* precursor (Clark,

1992). Batzle and Wang (1992) also considered the case of live oils and provided additional relationships to eqns [53] and [54] for such cases. They suggested, however, that additional work is necessary.

As noted by Clark (1992), the best way to obtain appropriate information is to directly measure sound speeds in the oils themselves. This approach can provide a great deal of thermodynamic information but it remains quite rare. Wang and coworkers carried out numerous measurements of the wave speeds in oils at ultrasonic frequencies and observed their variation with pressure and temperature (Wang and Nur, 1991; Wang et al., 1990). Daridon et al. (1998) conducted a series of ultrasonic measurements on a suite of hydrocarbons. Han and Batzle (2000) described a series of measurements on natural oils and provided a straightforward way to model the behavior. Oakley et al. (2003a,b) provided an extensive review of the literature of ultrasonic measurements in organic fluids in general and expressions for the speed of sound in 68 different pure organic fluids as a function of pressure.

Many workers have shown that the sound speeds in hydrocarbons are significantly affected by pressure and temperature. The experimental results of Daridon et al. (1998) on a light oil condensate, a medium weight hyperbaric oil, and a heavy natural hydrocarbon oil are shown in **Figure 13**. Generally, the denser the oil, the greater the wave speed. Wave speed also increases with pressure and decreases with temperature. However, it is important to note that these experiments were conducted at ultrasonic frequencies, and it is still not clear whether the values of K_f derived from them would reflect the adiabatic value at lower seismic frequencies; more work is required on this topic as very heavy oils have been shown to display a substantial dispersion both *in situ* (Schmitt, 1999) and in the laboratory (Behura et al., 2007).

Mixtures of CO_2 and H_2O are also of great importance for understanding chemical processes within the Earth. The growth of the geologic sequestration of greenhouse gases has further accelerated the need for knowledge of this system's behavior and its influence on the seismic properties of rocks

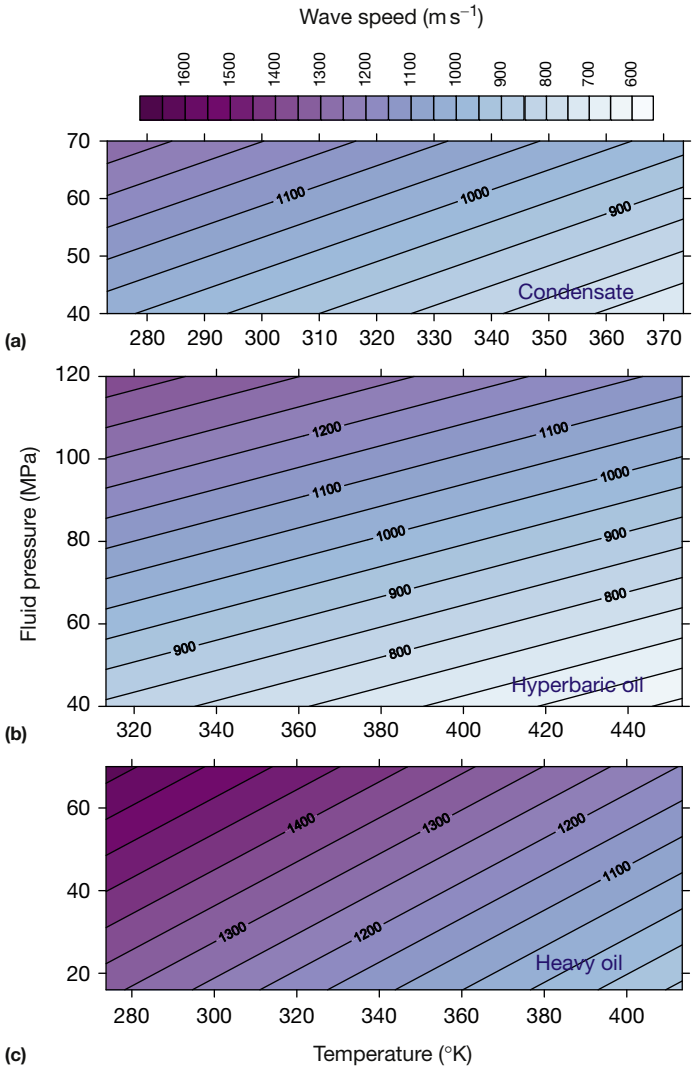


Figure 13 Sound wave speeds in (a) a light oil condensate, (b) a medium weight hyperbaric oil, and (c) a heavy undersaturated oil as measured by Daridon et al. (1998) with pressure and temperature.

in situ. Again, a large literature on the CO_2–H_2O system exists and Hou et al. (2013) gave an up-to-date listing of the many *PVTx* studies. The global phase diagram is described by a number of authors (Evelein et al., 1976; Spycher et al., 2003; Takenouchi and Kennedy, 1964; Wendland et al., 1999).

The limits on the solubility of CO_2 in H_2O complicate the phase diagrams (Duan and Sun, 2003) relative to those of the hydrocarbon mixtures as shown in **Figure 12** because, in addition to the vapor–liquid (VL) equilibrium curves, one must also consider the vapor–liquid–liquid (VLL) equilibrium where vapor coexists with H_2O-rich and CO_2-rich liquids. Some of this behavior is illustrated in **Figure 14**. Two compositionally dependent critical lines are evident although the one near the boiling line for pure CO_2 is small. The second critical line extends to higher pressures from the critical point of H_2O, but it does not continue to the critical point of CO_2 as was seen for the hydrocarbon mixtures in **Figure 12**. Instead, this curve bends backward to increasing temperatures with pressure, and it would be expected to eventually intersect the solid phase line at much higher pressures that are above the limits of current experiments (Evelein et al., 1976). This type of phase diagram is denoted class III (Gallagher et al., 1993; Gray et al., 2011; Vankonynenburg and Scott, 1980).

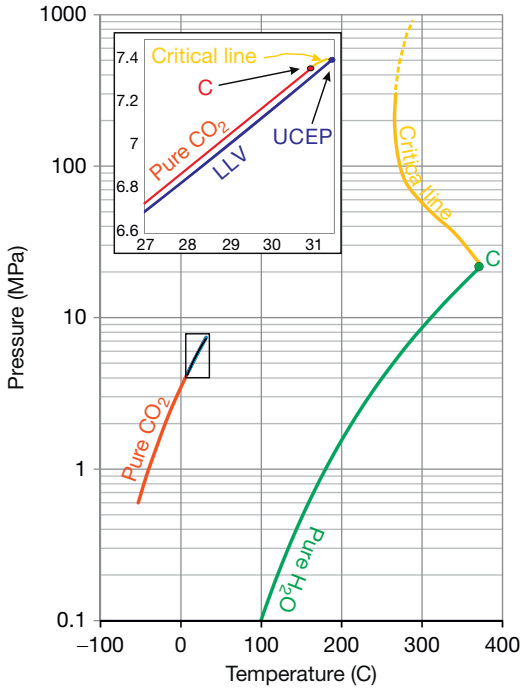


Figure 14 Some characteristics of the class III phase diagram of the CO_2–H_2O mixture system (Gray et al., 2011; Vankonynenburg and Scott, 1980) as projected onto the *P*–*T* plane. Inset is an expanded view of the area in *P*–*T* space near CO_2's critical point. Red and green lines are the boiling lines for single-component CO_2 and H_2O, respectively, and are the same as shown in **Figure 5(a)** and **5(b)**. These end at the respective critical points, also denoted as red or green C. Two composition-dependent critical lines exist. A short one connects CO_2's critical point to the upper critical end point (UCEP) at 7.411 MPa and 31.48 °C (Diamond and Akinfiev, 2003) for the liquid–liquid–vapor (LLV) three-phase line. This line is the locus of points where a H_2O-rich liquid, a CO_2-rich liquid, and a CO_2-rich vapor coexist with vapor existing to the right of this line. A second critical line extends to higher pressures from the critical point of water. Data for the LLV were determined by Wendland et al. (1999) while the second critical line is estimated from Takenouchi and Kennedy (1964). See also Gallagher et al. (1993).

Figure 14 highlights only the phase behavior of CO_2–H_2O mixtures. However, in order to calculate seismic properties, the density and bulk modulus of the mixture are required. Obtaining this information is challenging for such a complex system. Although a global model of the EOS for CO_2–H_2O mixtures has not yet been realized, workers have been able to provide EOS information over certain regions of *PVTx* space. For example, Duan and colleagues have developed models for the CO_2–H_2O gas phase from 0 to 28 MPa and from 323 to 645 °K (Duan et al., 2008) and vapor–liquid phase above 523 °K (Mao et al., 2009). Gallagher et al. (1993) provided estimates of behavior from 400 to 1000°K and to 100 MPa. However, a good deal of work still is required to fully understand this system (Hu et al., 2007). The use of such data in fluid substitution modeling of seismic responses may be further complicated by the fact that in practice, there will also be numerous additional impurities in the injected gas streams (Ziabakhsh-Ganji and Kooi, 2012).

11.03.5 The Rock Frame

11.03.5.1 Essential Characteristics

In order to properly understand the seismic properties of a real and usually fluid-saturated rock at depth in the Earth, one must also know the elasticity of the rock's skeleton or frame. The frame is an assemblage of a number of solid minerals. This construction, free of fluids or dry, will have its own elastic properties K_d and μ_d and density ρ_d. The subscript 'd' usually denotes the 'dry' or unsaturated state for the frame. This can also mean in certain contexts 'drained,' which is the state of a saturated sample after sufficient time such that the pore fluid pressure has equilibrated after a deformation. In the literature, the 'dry' and 'drained' conditions are usually taken to give the same static bulk modulus. The situation is different for wave propagation as fluids may not have time to move relative to the frame. In the earlier literature, the reader needs to take care when the term 'skeleton' is discussed as this has variously been used to describe either the mineral or the frame. The frame is characterized by its constituent minerals that will have their own solid moduli K_m and μ_m and density ρ_m, its porosity ϕ, and its permeability κ.

The fluid directly affects the overall density of the rock, ρ, through eqn [21] that is often simplified to

$$\rho = \phi\rho_f + (1-\phi)\rho_m \qquad [55]$$

where ϕ is the porosity (the ratio of the void space volume to the envelope or total volume of the material) and ρ_f and ρ_m are the fluid and the solid densities, respectively. When more than one fluid resides in the pore space, eqn [55] is further modified, particularly in the applied literature, to saturations S_j for each of the n immiscible free fluid components such that

$$\rho = \phi\left(\sum_{j=1}^{n} S_j\rho_j^f\right) + (1-\phi)\rho_m \qquad [56]$$

with

$$1 = \sum_{j=1}^{n} S_j \qquad [57]$$

and where ρ_f^f is the respective fluid component density. Usually, the saturations are assigned more directly to the different fluid phases such as S_w, S_o, and S_g for water, oil, and gas, respectively.

With these definitions out of the way, we proceed in this section to discuss the elastic properties of the rock frame. These frame properties are key to understanding the elasticity, and hence the wave propagation, through rocks. Again, there is a large literature on this topic as the study of wave propagation in porous fluid-saturated media is important in many disciplines.

11.03.5.2 The Pore-Free Solid Portion

In this section, we are interested in the effective elastic properties of the solid frame by itself. Rocks are most commonly a mixture of a number of minerals, and often, knowledge of the effective bulk K_m and shear μ_m moduli for the solid mineral portion is a first prerequisite to the more sophisticated calculations incorporating porosity and fluids. To obtain these, workers often assume that on average, the rock's minerals are randomly oriented such that the effective material is elastically isotropic. The simplest approaches to calculating the effective mineral moduli of the pore-free composite were derived by Voigt (1887) and by Reuss (1929a) who, respectively, assumed constant strain or stress within each of the components. Consequently, the Voigt and Reuss formulations are mixing rules for the stiffnesses or the compliances, respectively, such that for N isotropic components,

$$K_V = \sum_{i=1}^{N} \phi_i K_i \geq K_m \geq \left[\sum_{i=1}^{N} \phi_i / K_i\right]^{-1} = K_R \qquad [58]$$

and

$$\mu_V = \sum_{i=1}^{N} \phi_i \mu_i \geq \mu_m \geq \left[\sum_{i=1}^{N} \phi_i / \mu_i\right]^{-1} = \mu_R \qquad [59]$$

In eqns [58] and [59], the values of K_i and μ_i could be taken from **Table 2** or more extensive compilations (Bass, 1995).

Hill (1952) demonstrated that the Voigt and Reuss values bound the composite's effective bulk K_m and shear μ_m moduli. Various authors (see Watt et al. (1976)) have argued over which of the bounds or averages provide the best values. In practice, Hill's simple arithmetic mean of the bounds $K_{VRH} = [K_R + K_V]/2$ and $\mu_{VRH} = [\mu_R + \mu_V]/2$ is most usually calculated to give K_m and μ_m as done similarly in **Section 11.03.3.2** for isotropic composites of polycrystals made of a single mineral phase. The difference here is that the bounds are assumed to be mixtures of isotropic materials.

Hashin and Shtrikman's (1962) approach may also be applied to estimating the composite's elastic moduli according to

$$K_{HS}^{+} = K_1 + \frac{\phi_2}{(K_2 - K_1)^{-1} + \phi_1\left(K_1 + \frac{4}{3}\mu_1\right)^{-1}} \geq K_m$$

$$\geq K_2 + \frac{\phi_1}{(K_1 - K_2)^{-1} + \phi_2\left(K_2 + \frac{4}{3}\mu_2\right)} = K_{HS}^{-} \qquad [60]$$

and

$$\mu_{HS}^{+} = \mu_1 + \frac{\phi_2}{(\mu_2 - \mu_1)^{-1} + \frac{2\phi_1(K_1 + 2\mu_1)}{5\mu_1\left(K_1 + \frac{4}{3}\mu_1\right)}} \geq \mu_m$$

$$\geq \mu_2 + \frac{\phi_1}{(\mu_1 - \mu_2)^{-1} + \frac{2\phi_2(K_2 + 2\mu_2)}{5\mu_2\left(K_2 + \frac{4}{3}\mu_2\right)}} = \mu_{HS}^{-} \qquad [61]$$

Equations [60] and [61] are valid for only two mineral components, but Hashin and Shtrikman (1963) also provided the methodology to carry out the calculations for N components (Berryman, 1995; Watt et al., 1976).

Calculations using eqns [58] through [61] are carried out for a hypothetical mixture of quartz and calcite for purposes of illustration in **Figure 15**. Such a quartz–calcite mixture is not expected to exist naturally, but these two minerals were chosen only because of the large differences in their elastic moduli and because of the unusual differences between K_m and μ_m. Equation [59] gives similar values of the shear moduli with the largest deviation between μ_V and μ_R being $< 3\%$ at $\phi_{quartz} = 0.5$. In contrast, the Voigt and Reuss bulk moduli from eqn [58] differ by nearly 11% at the same quartz proportion. In contrast, the Hashin–Shtrikman bounds for both moduli of eqns [60] and [61] differ by only a fraction of a percent. This translates into tighter differences for the wave speeds that vary by at most by 3.5% and 1.4% for V_P and V_S, respectively. Consequently, eqns [58] through [61] are widely used to estimate K_m and μ_m for the rock, and there appears to be good experimental evidence of their applicability to pore-free metal composites (Hashin and Shtrikman, 1963; Umekawa and Sherby, 1966) and to an assortment of rocks (Brace, 1965; Ji et al., 2002) subject to high pressures.

One caveat is that in many rocks, the minerals have a preferential crystallographic alignment due to sedimentary deposition or metamorphic deformation. In such cases, the minerals are no longer crystallographically randomly oriented; and their intrinsic anisotropy affects the overall anisotropy of the composite. These preferential mineral alignments are variously called lattice-preferred orientations (LPO) or crystallographic-preferred orientations (CPO) in the literature. This requires that the various mixing theories be modified to account for the crystal symmetries and the statistics of their orientations with respect to that of the material.

There is insufficient room to go into details of the procedures used and only a brief listing of the relevant literature is given. Different averaging techniques were described by a variety of workers (Babuska, 1972; Bunge, 1974; Crosson and Lin, 1971; Kumazawa, 1969; Thomsen, 1972). Mainprice (1990) provided computer programs that utilized these ideas and applied it to a textured plagioclase rock.

Noting that the results from the Voigt and Ruess averages are not invertible to each other (i.e., eqn [13] does not hold), Matthies and Humbert (1993) developed the geometric mean average that was then used by Mainprice and Humbert (1994) on polycrystals of feldspar and biotite and by Cholach and Schmitt (2006) to explore the effects of the strength of phyllosilicate orientations on the seismic anisotropy of shales and schists. The ability to quantitatively obtain mineralogical orientation distribution functions has grown greatly in the last

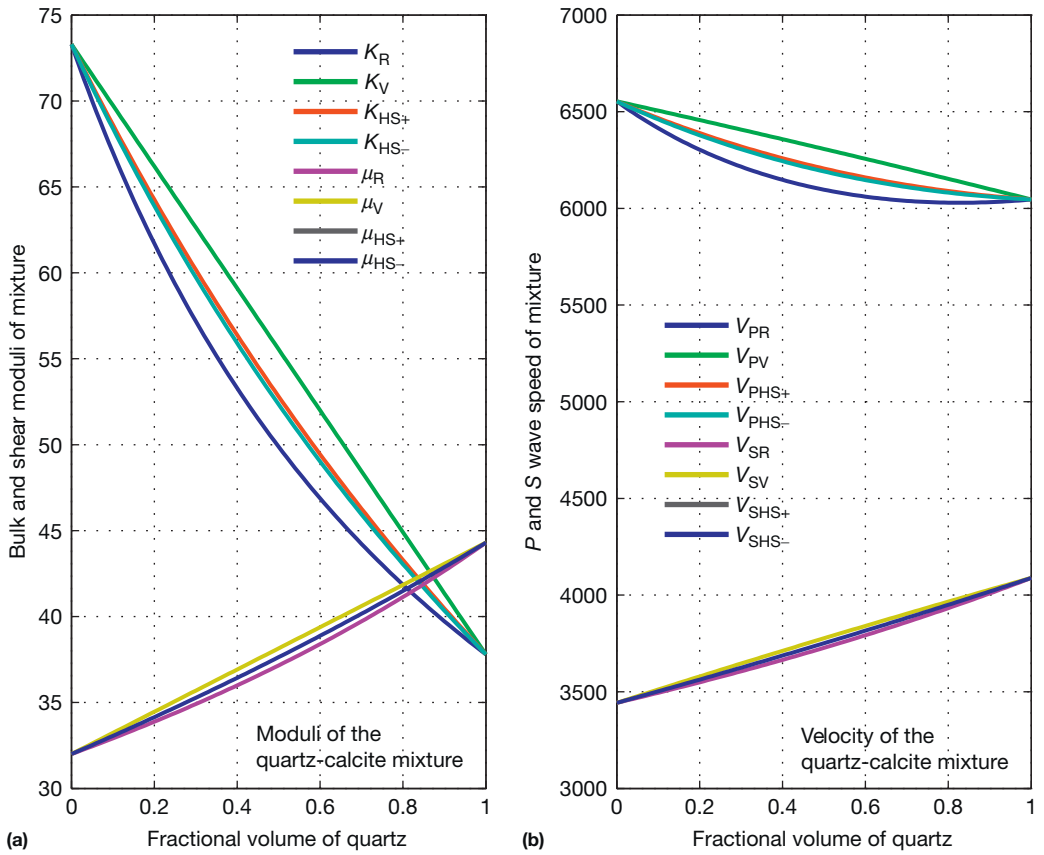


Figure 15 Illustration of the predictions of the (a) Voigt, Reuss, and Hashin–Shtrikman bounds and the resulting (b) V_P and V_S for a hypothetical mixture of calcite and quartz as a function of the volume fraction of quartz.

decade with developments in x-ray and neutron scattering and electron backscatter diffraction techniques. This is motivating a growth in the use of the averaging techniques that have now been applied to numerous different rock types (e.g., Almqvist et al., 2010; Kanitpanyacharoen et al., 2011; Wenk et al., 2007, 2012), although often the model results disagree with corresponding observations because the averaging models cannot account for porosity. Development of modeling procedures still continues (Man and Huang, 2011; Morris, 2006).

11.03.5.3 Influence of Porosity

The rock's porosity ϕ has a large influence on the elastic properties of a material. Walsh et al.'s (1965) classic experiments still remain of interest in this context. They constructed a series of glass 'foams' of differing porosities by heating packs of glass beads. They then measured the linear compressibility of these samples with strain gauges during hydrostatic compression; unfortunately, this only provides a measure of the K_d as indicated by the filled squares in **Figure 16(a)**. Regardless, these data illustrate the rapid decline in K_d with ϕ.

It is worthwhile to compare these observations to the Voigt, Ruess, and Hashin–Shtrikman bounds described above, the expressions developed for a solid composed of a 'mineral' filled with spherical pores by MacKenzie (Li and Zhang, 2011; Mackenzie, 1950; Yoshimura et al., 2007) and those of Kuster and Toksöz (1974a,b).

Mackenzie's expressions take the form

$$K_{MK} = \frac{K_m}{\left[1 + \frac{3\phi(1-\nu_m)}{2(1-2\nu_m)(1-\phi)}\right]} \quad [62]$$

and

$$\mu_{MK} = \frac{\mu_m(1-\phi)}{\left[1 + \frac{\phi(12+6K_m/\mu_m)}{(8+9K_m/\mu_m)}\right]} \quad [63]$$

Poisson's ratio for the solid portion ν_m may be calculated from the pore-free mineral moduli K_m and μ_m using a standard elastic relation (Birch, 1961)

$$\nu_m = \frac{1}{2}\frac{3K_m - 2\mu_m}{3K_m + \mu_m} \quad [64]$$

The corresponding formulas for spherical inclusions with moduli K_i and μ_i and volume fraction, ϕ, according to Kuster and Toksöz (1974a) are

$$(K_{KT} - K_m)\frac{(K_m + 4\mu_m/3)}{(K_{KT} + 4\mu_m/3)} = \phi_i(K_i - K_m)\frac{(K_m + 4\mu_m/3)}{(K_i + 4\mu_m/3)} \quad [65]$$

and

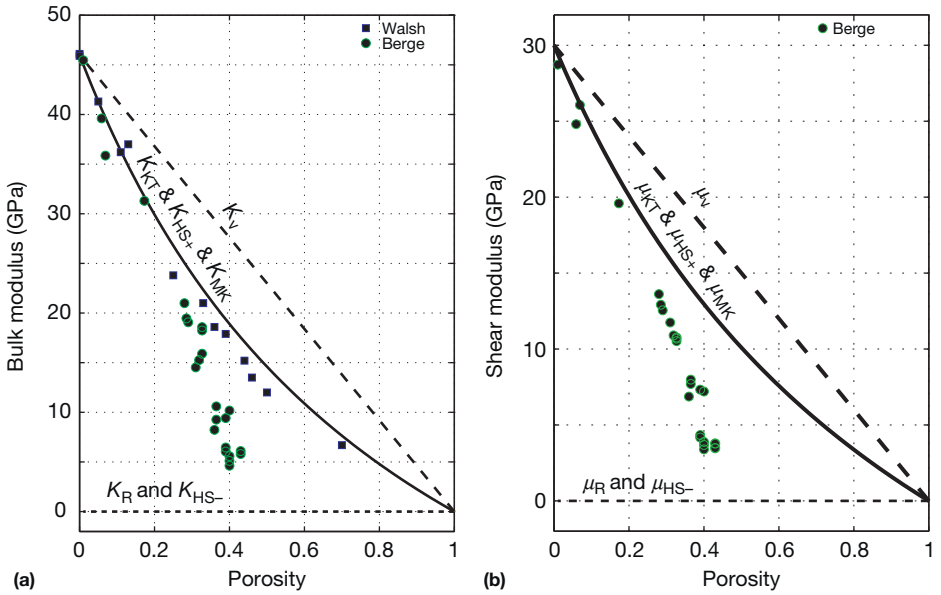


Figure 16 Comparison of experimentally observed (a) bulk modulus and (b) shear modulus versus air-filled porosity in sintered porous glass composites from Walsh (1965) and Berge et al. (1995). Subscripts on the moduli are R, Ruess bound; V, Voigt bound; HS− and HS+, Hashin–Shtrikman bounds; KT, Kuster–Toksöz; and MK, Mackenzie.

$$\left(\mu^{KT}-\mu_m\right)\frac{(\mu_m+\zeta_m)}{(\mu^{KT}+\zeta_m)}=\phi_i(\mu_i-\mu_m)\frac{(\mu_m+\zeta_m)}{(\mu_i+\zeta_m)} \qquad [66]$$

where

$$\zeta_m=\frac{\mu_m(9K_m+8\mu_m)}{6(K_m+2\mu_m)} \qquad [67]$$

The moduli for the inclusions are taken to be zero for the calculations shown in **Figure 16**. For this particular case (and also for cases in which the spherical voids are filled with a fluid for which $\mu_i=0$), the Reuss and the lower Hashin–Shtrikman bounds coincide. Further, the Kuster–Toksöz and the upper Hashin–Shtrikman bound values also are the same, and these are also close to those calculated using Mackenzie's equations.

One important observation from **Figure 16** is that unlike the calculations for a nonporous mixture of solids (**Figure 15**), the Voigt–Reuss bounds are less useful in predicting the moduli for porous materials. One reason for this is the large divergence between the vanishing moduli in the empty pore. The Hashin–Shtrikman bounds, too, have a wide variance although the upper bounds appear to match Walsh et al.'s (1965) bulk modulus observations reasonably well.

There is some scatter in Walsh et al.'s (1965) experimental results in **Figure 16(a)**, but they do trend close to but slightly lower than the $K_{KT}-K_{HS+}-K_{MK}$ prediction line. The reason for this most likely lies in the fact that both the Kuster–Toksöz and Mackenzie formulations assume 'dilute' amounts of spherical voids that are sufficiently removed from one another such that the interactions between their concentrated stress and strains may be ignored. Interaction of the elastic fields from individual spheres cannot be avoided once the spheres become sufficiently proximate.

Numerous authors have attempted to overcome these limitations by addressing such interactions by two main groups of approaches: that of an effective matrix or that of an effective field (Carvalho and Labuz, 1996; Mavko et al., 2003). The effective matrix approaches are often further subdivided either into 'self-consistent' methods (Budiansky and O'Connell, 1976; Hill, 1965; Wu, 1966) that consider the solids and pores as a whole or by 'differential effective medium' methods in which pores are incrementally added to a matrix in order to update its modulus, and this revised modulus is then that of the new matrix for the next iteration (e.g., Berryman et al., 2002; David and Zimmerman, 2011; Kachanov, 1980; Li and Zhang, 2011; Norris, 1985; Zimmerman, 1984). The effective field methods developed by Mori and Tanaka (1973) place a pore in the nonporous solid to which a stress field is then applied. The advantage of this approach is that more complex stress fields can be applied within the material accounting directly for the existence of the holes. These techniques have been widely applied in the composite materials literature but have not seen as much application in rock physics (Sayers and Kachanov, 1995), although Kachanov et al. (1994) suggested that the Mori–Tanaka approach may be superior. Berryman and Berge (1996) noted the exclusivity of the use of the Mori–Tanaka approach in the engineering literature and contrasted it with the Kuster–Toksöz model employed in geophysics indicating that both are likely only valid for $\phi<0.3$. Yan et al. (2011) gave a recent example of the use of the Mori–Tanaka approach in estimating the frame moduli of sandstones.

The literature on modeling of rock properties using the various developments presented earlier in the text is large and only a small view of it is possible here. Mavko et al. (2003) provided some starting points for further investigations.

The porosity-dependent variations in moduli transfer through to the wave speeds according to eqns [1] and [2]. An illustration of this effect in a suite of carbonates, primarily clean limestones, shows the significant changes in both V_P and V_S with increasing porosity. Both wave speeds decrease

by upward of 50% over the range of ϕ measured. This is despite the fact that the bulk density, ρ, which is in the denominator of eqns [1] and [2], also decreases as ϕ increases according to eqn [55]. Cleary, porosity is an important factor in determining a material's seismic wave speeds.

11.03.5.4 Influence of Crack-Like Porosity

As just noted, aside from the constituent mineralogy, the most important factor affecting the elastic properties of the frame is ϕ. The calculations shown in **Figure 16** are based on spherical pores, but most pores will not be shaped so simply; and the geometry of the pores needs also to be considered. It is useful to return to **Figures 16** and **17** to see some of such effects. Berge et al. (1995) measured V_P and V_S on sandstone analogs of lightly sintered glass beads and, from these values, calculated the dynamic moduli shown in **Figure 16**. Their observed values substantially deviate from those obtained in Walsh et al.'s (1965) measurements and from those predicted by Mackenzie's (1950) derivation of eqn [62]. Bakhorji's (2009) measurements on carbonates, too, show considerable scatter of up to 1700 m s^{-1} for limestones near a given ϕ. Other factors that simplify porosity and mineralogy must be influencing these observations.

Much of this can be explained by considering the influence of the pore shape. At the risk of oversimplifying the problem, pore shapes can broadly be placed into two categories: equant or crack-like. Much like a Roman arch, large aspect ratio equant pores are stiff; they do not significantly deform under the application of a stress. In contrast, small aspect ratio crack-like pores are easily compressed and can close under even modest stresses.

An important consequence of this is that the rock elastic moduli are generally pressure-dependent, a fact that has been known since the very first elastic measurements on rock to high pressures by Adams and Williamson (1923) (**Figure 18(a)**). They were surprised that such effects were seen in rocks with porosities $<1\%$. They theoretically tested a number of possible

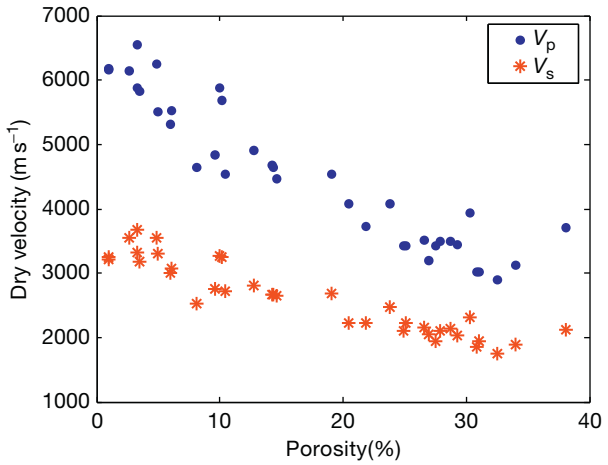


Figure 17 Illustration of the effects of porosity on V_P (blue filled circles) and V_S (red stars) for a suite of carbonates from the Arab D formation of Saudi Arabia. Experimental errors are approximately the size of the symbols. Unpublished ultrasonic measurements from Bakhorji (2009).

hypotheses and finally inferred that this nonlinear behavior must be due to the existence of crack-like pores of small aspect ratio.

Heuristically, this behavior stems from the large compressibility of a crack-like pore perpendicular to its plane. A normal stress applied across the crack face pushes the surfaces towards one another. As this stress increases, the effective crack shortens and eventually closes at pressure P_c. For a crack with an elliptical cross-section of aspect ratio χ (ratio of the minor to major axes), this is (Walsh, 1965)

$$P_c = \frac{\pi E_m \chi}{4\left(1-\nu_m^2\right)} \qquad [68]$$

where E_m and ν_m are the solid mineral Young's modulus and Poisson's ratio, respectively. Cracks with small aspect ratios χ are easily closed. For example, a hypothetical elliptical crack with $\chi \sim 10^{-3}$ residing in quartz matrix closes by application of <2 atmospheres of pressure ($\sim$0.2 MPa). In contrast, a spherical 'crack' with $\chi = 1$ would according to eqn [68] not close until 172 GPa. Of course, the material would crush long before such a pressure could be reached, but these calculations serve to illustrate how easily crack-like porosity can be closed relative to more equant porosity.

Consider subjecting a material containing cracks with a distribution of different χ's to an increasing confining pressure. At low pressures, those cracks with small χ first close. As the pressure continues to increase, cracks with progressively larger χ will close. Once closed, a crack no longer influences the overall elasticity and essentially disappears with the result that the rock becomes less compressible. This is illustrated in **Figure 18(b)**, which is a cartoon of the expected strains on a cracked rock sample. As pressure increases, the observed strain (heavy dark line) is initially nonlinear. As the cracks progressively close, the material stiffens and the strain–stress curve becomes less steep. This continues until all of the crack porosity is closed at which point the strain–stress curve becomes linear and parallel to that expected for the pore-free solid (thick dashed line). The observed strain is a combination of the strains due to closure of the cracks (thick dashed line) and that of the solid. This can be seen in the real strains observed on a dolomite rock sample (**Figure 18(c)**). These arguments are perhaps a bit oversimplified (Stroisz and Fjaer, 2013), but regardless at lower confining stresses, rocks are generally nonlinearly elastic materials, that is, their moduli and subsequently their seismic wave speeds depend on confining pressure. One may be able to ignore this in the deep crust and the Earth's mantle, but it appears to be an important factor in seismic investigations nearer the Earth's surface (Crampin and Peacock, 2008; Schijns et al., 2012). Recognizing this fact is key, for example, in properly interpreting time-lapse seismic observations from reservoirs subject to varying states of effective stress.

As noted, such nonlinear pressure dependencies have been noted since Adams and Williamson (1923). This pressure dependence of the elastic moduli translates into a pressure dependence of the waves speeds. Such effects have been observed in nearly all rock types and an incomplete listing of contributions where such nonlinear effects have been observed includes those for igneous and metamorphic hard rocks (e.g., Birch, 1960, 1961; Blake et al., 2013; Cheng and Toksöz, 1979; Cholach et al., 2005; Christensen

and Stanley, 2003; Goddard, 1990; Kern, 1982; Lyakhovsky et al., 1997; Todd and Simmons, 1972), sandstones (e.g., Christensen and Wang, 1985; Gomez et al., 2010; Jones and Nur, 1983; Khazanehdari and McCann, 2005; Khazanehdari and Sothcott, 2003; Lo et al., 1986; Prasad and Manghnani, 1997; Sayers, 2002; Sayers et al., 1990; Smith et al., 2010; Wyllie et al., 1958; Xu et al., 2006), carbonates (e.g., Alam et al., 2012; Azeemuddin et al., 2001; Fabricius et al., 2008; Melendez Martinez and Schmitt, 2013), and mudstones (Freund, 1992; Kwon et al., 2001; Sayers, 1999). For example, the wave speeds in a highly cracked sandstone (**Figure 19(b)**) from the Cadotte formation in Alberta vary greatly even over the relatively modest range of confining pressures to 60 MPa (He, 2006). It is interesting to contrast this rock's behavior with that for a comparison brass sample that over this pressure range has a wave speed that increases slightly due to the pressure-dependent increase in the intrinsic crystal moduli as described in the section on minerals in the preceding text.

Many workers have focused on different ways to describe this nonlinearity including the use of curve fitting (Carcione and Tinivella, 2001; Eberhart-Phillips et al., 1989; Freund, 1992; Khaksar et al., 1999; Kirstetter et al., 2006; Prasad and Manghnani, 1997; Prikryl et al., 2005; Zimmerman, 1985), Birch–Murnaghan equations of state (Birch, 1961), differential

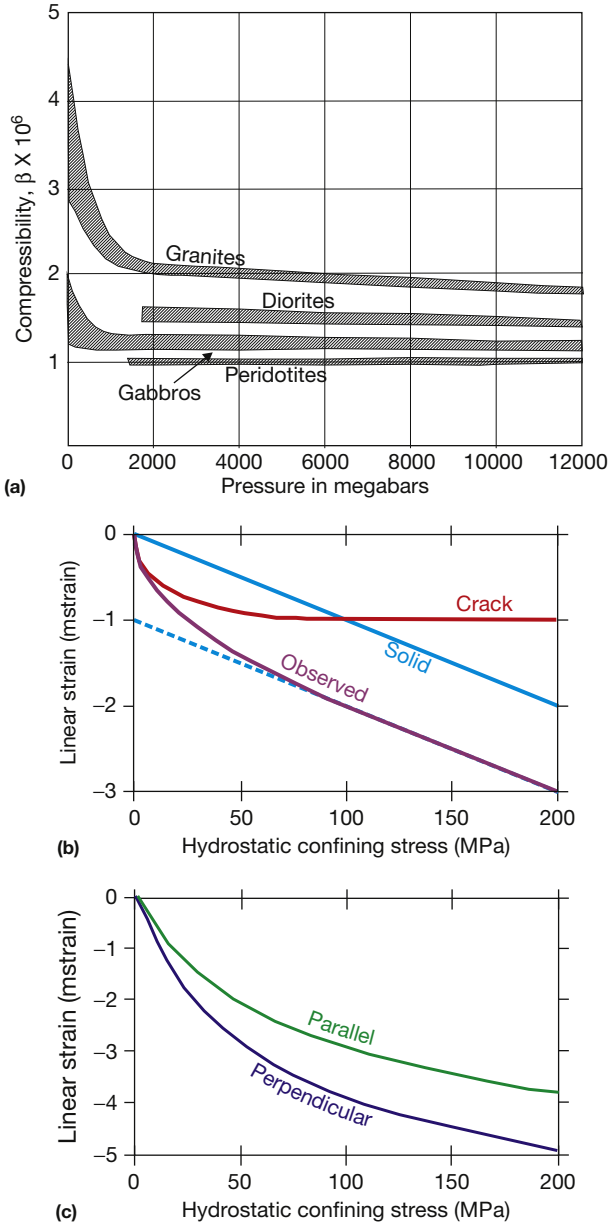


Figure 18 Effect of cracks on the elastic properties of rocks. (a) Early observations of compressibility of a series of rocks, as shown in the redrafting of Figure 7 of Adams and Williamson (1923). The gray areas represent the spread of the moduli they expected due to mineralogical differences within a given geologically defined rock type. Note that their definition of a megabar = 1.0197 kg cm^{-2} = 0.9869 atm ~ 0.1 MPa and should not be confused with the modern Mbar = 100 GPa. (b) Illustration of the partitioning of the total observed strain into components due to the cracks and the minerals as a function of pressure (redrafted from Schmitt and Li, 1995). (c) Observed strains parallel and perpendicular to the sedimentary layering in a cracked dolomite as a function of hydrostatic confining pressure. (data from Schmitt and Li, 1995).

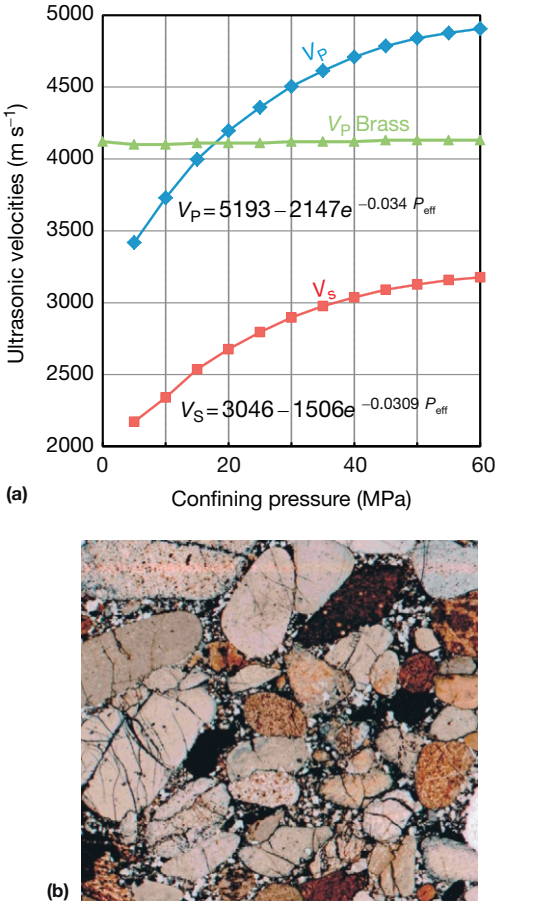


Figure 19 Example of the effects of pressure on wave speeds in a cracked rock. (a) Ultrasonically determined V_P (blue diamonds) and V_S (red squares) as a function of hydrostatic confining pressure. V_P for nonporous brass (green triangles) is shown for comparison. The fits to eqn [70] are shown. (b) Photomicrograph (2 × 2 cm) under transmitted light of the Cadotte sandstone showing its highly cracked nature. Despite differences in the color of the pebbles, they are nearly pure microcrystalline quartz. Unpublished data (for sample SB009 from 2457.7 m depth) and image from He (2006).

approaches (Ciz and Shapiro, 2009), crack damage (Lyakhovsky et al., 1997), and third-order elastic moduli (Payan et al., 2009; Sayers and Kachanov, 1995; Sinha and Kostek, 1996). In describing the pressure dependence of a velocity in such rocks, often workers will use an empirical expression of the form

$$V_i = A + CP_{\text{eff}} - Be^{-DP_{\text{eff}}} \quad [69]$$

where V_i can be either V_P or V_S, P_{eff} is the effective pressure (to be described shortly) applied to the material, and A, B, C, and D are simply parametric fitting parameters. C is often left as zero as it can result in unreasonable values at elevated pressures (Khaksar et al., 1999). This is convenient when the only information available might be laboratory measurements on the core samples. While eqn [69] is essentially devoid of any physics, it does well describe the shapes of the curves shown in **Figure 19** differing from the observed values by <15 m s^{-1} (He, 2006).

A number of approaches have been employed to more fundamentally explain the effects of cracks on elastic moduli and velocities. The already mentioned self-consistent approach (Budiansky and O'Connell, 1976; O'Connell and Budiansky, 1974) remains popular because of its relative ease of application. Consider the case of flat circular cracks of radius a and with an elliptical cross-section of minor axis c and with $a \gg c$. The crack density parameter is $\epsilon \equiv N < a^3 >$ where N is the number of cracks per unit volume. It can only be related to the crack porosity ϕ_c if one assigns values to a and c so that $v_c = 4\pi a^2 c/3$ if all the cracks have the same dimensions whereupon $\phi_c = Nv_c$. Once ϵ is set, the effective Poisson's ratio of the cracked solid v_d must first be found from

$$\epsilon = \frac{45}{16}\frac{(v_m - v_d)(2 - v_d)}{(1 - v_d^2)(10v_m - 3v_m v_d - v_d)} \quad [70]$$

whereupon the moduli

$$K_d = K_m\left[1 - \frac{16}{9}\left(\frac{1 - v_d^2}{1 - 2v_d}\right)\epsilon\right] \quad [71]$$

and

$$\mu_d = \mu_m\left[1 - \frac{32}{45}\frac{(1 - v_d)(5 - v_d)}{(2 - v_d)}\epsilon\right] \quad [72]$$

are easily calculated. For purposes of illustration, this self-consistent approach is applied by placing cracks with major axis diameter of $a = 1$ mm for three different and small aspect ratios $\alpha = c/a$ in **Figure 20**. We choose to plot the moduli and wave speeds as function of the crack porosity ϕ_c because the physical meaning of the crack density parameter ϵ is less intuitive. Examination of **Figure 20** shows that smaller aspect ratio cracks have a disproportionate influence on the overall elastic moduli and velocities. Further, even a vanishingly small crack porosity of $<1\%$ has a substantially larger effect than the equivalent more equant porosity. This is particularly apparent when the theoretical results for spherical pores of **Figure 16** are compared to the equivalent crack porosities in **Figure 20**.

While we have used here the O'Connell and Budiansky (1974) self-consistent forms to illustrate the large influence of crack-like porosity on rock properties, it must be noted that there are also many competing models. Again, geophysicists

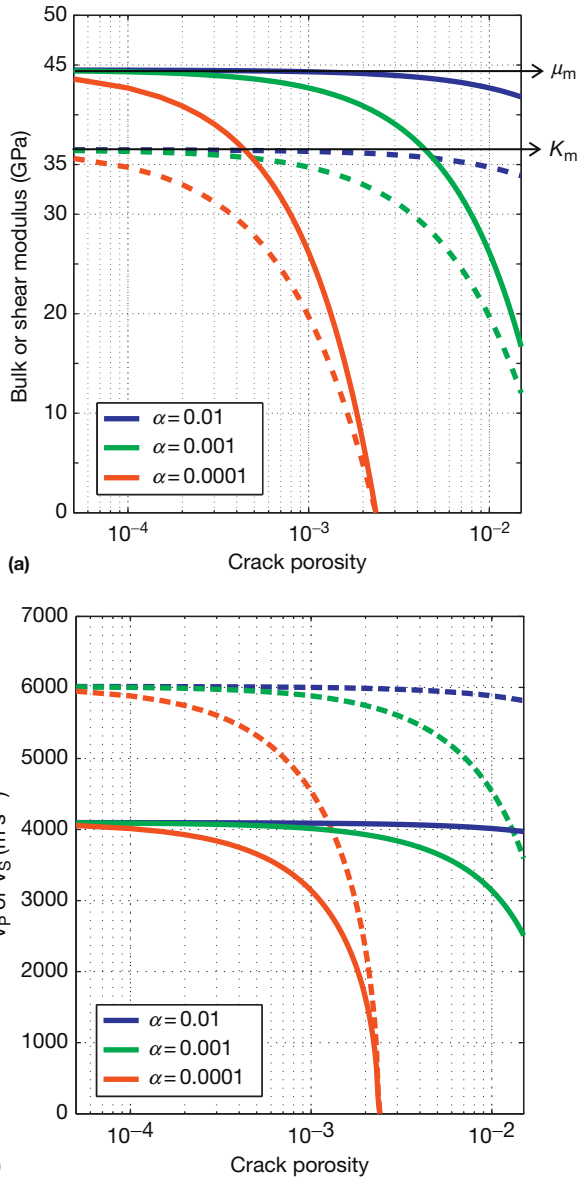


Figure 20 Illustration of the influence of crack porosity for elliptical pores with $a = 1$ mm and with aspect ratios $c/a = 0.01$ (blue lines), 0.001 (green lines), and 0.0001 (red lines) in a quartz solid on (a) the elastic moduli K_d (solid lines) and μ_d (dashed lines) and (b) the wave speeds V_P (dashed lines) and V_S (solid lines) according to self-consistent models.

will often employ the Kuster and Toksöz (1974a) model adjusted to employ small α ellipsoids. Horii and Nemat-Nasser (1983) extended the self-consistent approach by including crack–crack interactions, crack closure, and frictional sliding along the planes of closed cracks to show that the loading history may be important. Berryman et al. (2002) developed a differential effective medium model to account for dry and fluid-filled cracks. Hudson (1981) used a mean field approach to account for the dynamic effects of both aligned and randomly oriented cracks, and an advantage of his formulation is that he can account for seismic attenuation. Mavko and Nur (1978) carried out an analysis that employed dislocation theory to solve for the deformations associated with a more realistic

crack geometry with tapered edges. More recently, Gao and Gibson (2012) developed a statistical asperity model to describe the influence of microcracks and provide a more extensive review of the recent literature than is possible here.

Numerous authors have attempted to use the shapes of either the strain curves (as in **Figure 18**) or wave speeds (as in **Figure 19**) to invert for various characteristics of the cracks and the distributions of their dimensions (e.g., Angus et al., 2009; Cheng and Toksöz, 1979; David and Zimmerman, 2012; Schubnel et al., 2006) under a variety of simplifying assumptions.

11.03.5.5 Pressure Dependence in Granular Materials

Cracked Earth materials are not the only ones with pressure-dependent frame properties. Granular materials, such as unconsolidated sands at low confining pressures, too, are also highly sensitive to applied confining pressures. This has been demonstrated by many workers in the laboratory (Bachrach and Avseth, 2008; Domenico, 1977; Goddard, 1990; Makse et al., 1999; Zimmer, 2003; Zimmer et al., 2007) or inferred from field measurements (Bachrach et al., 2000). In order to predict the moduli and wave speeds in such materials, theoretical developments (Digby, 1981; Walton, 1987) rely primarily on the 'Hertz–Mindlin' deformations (Johnson, 1987) at the contact between two mineral grains. Consider a grain pack with porosity ϕ within which each grain will on average touch R neighbors. Further, assume that both normal and shear tractions exist at each grain contact according to Walton's (1987) 'infinitely rough' grain model. The moduli of this grain pack as rewritten by Makse et al. (1999) are

$$K_{\mathrm{d}} = \frac{C_{\mathrm{n}}(\phi R)^{2/3}}{12\pi}\sqrt[3]{\frac{6\pi P}{C_{\mathrm{n}}}} \qquad [73]$$

and

$$\mu_{\mathrm{d}} = \frac{(C_{\mathrm{n}} + 3C_{\mathrm{t}}/2)(\phi R)^{2/3}}{20\pi}\sqrt[3]{\frac{6\pi P}{C_{\mathrm{n}}}} \qquad [74]$$

where

$$C_{\mathrm{n}} = \frac{4\mu_{\mathrm{m}}}{1-\nu_{\mathrm{m}}} \qquad [75]$$

and

$$C_{\mathrm{t}} = \frac{8\mu_{\mathrm{m}}}{2-\nu_{\mathrm{m}}} \qquad [76]$$

are factors that relate the forces at the contact points to the induced overlap between the grains. Walton also derived expressions for the case of 'smooth' spheres where any tangential forces at the grain contacts disappear. Makse et al. (1999) gained insight from discrete particle modeling (Cundall and Strack, 1979) to infer that the coordination number R is also pressure-dependent because of grain motions with a value empirically derived from their numerical modeling of

$$\langle R\rangle = 6 + \left(\frac{P}{0.06}\right)^{1/3} \qquad [77]$$

where in this expression, the confining pressure P must be in units of MPa. The cubic root dependence of the moduli with confining pressure in eqns [73] and [74] is perhaps the most interesting aspect of these equations. This is essentially due to the fact that the grain contacts become stiffer with pressure as their contact area expands. The moduli increase more rapidly in Makse et al.'s (1999) model because the number of grain contacts is also increasing with pressure. This cubic root for the moduli translates to a 1/6 root dependence for the wave speeds.

Although the theory illustrated in **Figure 21** captures some of the elements of the pressure dependence of the moduli, real measurements on granular materials give different moduli and

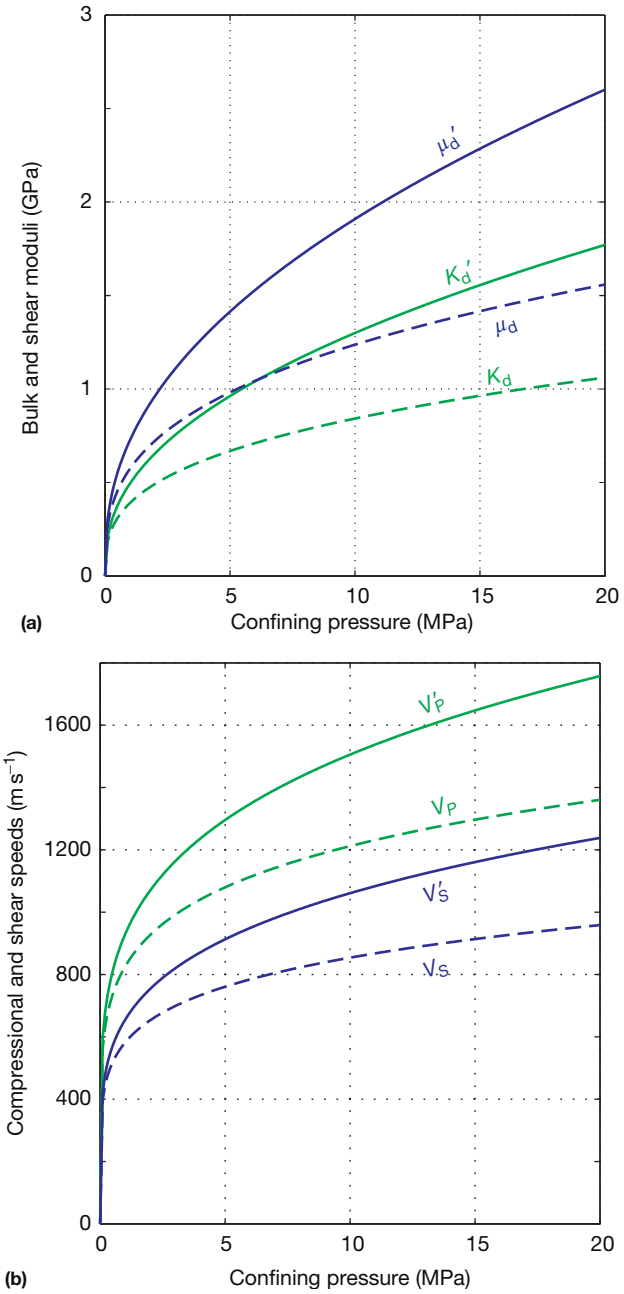


Figure 21 Illustration of the confining pressure dependence of (a) the dry frame moduli K_{d} (eqn [74], dashed green line) and μ_{d} (eqn [75], dashed blue line) according to Walton (1987) and (b) the corresponding compressional V_{P} (dashed green line) and shear V_{S} (dashed blue line) wave speeds of an unconsolidated pack of quartz spheres (K_{m} = 36.5 GPa, μ_{m} = 44.5 GPa, ρ_{m} = 2650 kg m^{-3}) with a porosity ϕ = 0.36 and coordination number R = 6. The corresponding values that include the pressure dependence of R (eqn [78]) are shown as solid lines and denoted by ′.

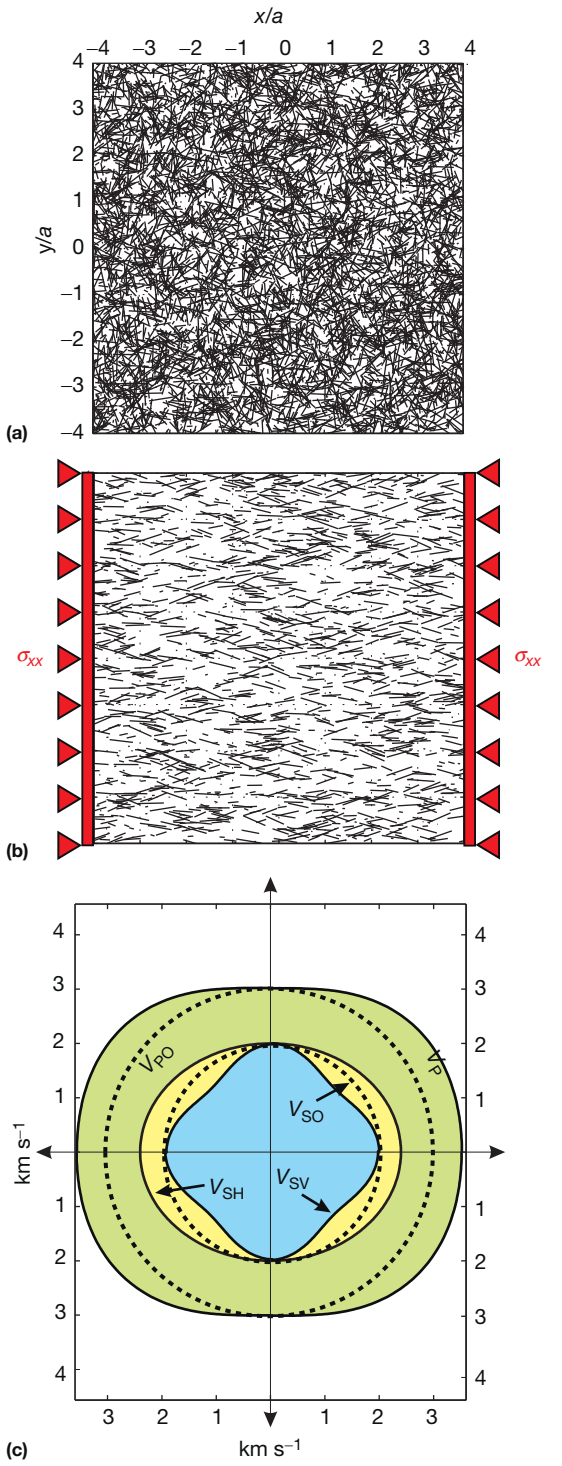


Figure 22 Two-dimensional description of crack-induced elastic anisotropy. (a) Initial cracked medium with random orientation of cracks. (b) Distribution of remaining open cracks after application of horizontal compression stress σ_{xx} (modified from Schmitt DR, Currie CA, and Zhang L (2012) Crustal stress determination from boreholes and rock cores: Fundamental principles. *Tectonophysics* 580: 1–26). (c) Hypothetical wave velocity surfaces before (circles outlined with dashed lines) and after (solid lines) application of σ_{xx}. V_{PO} and V_{SO} are the original isotropic compressional and shear wave speeds material of (a). V_P, V_{SV}, and V_{SH} are the induced anisotropic wave surfaces for the compressional, the vertically polarized shear (i.e., in the plane of the figure), and the horizontally polarized shear (i.e., perpendicular to the plane of the figure) wave modes, respectively. In (c), the wave speed is represented by the distance from the origin.

velocities (Hardin and Blandford, 1989). On the basis of an extensive series of ultrasonic wave speed measurements on a wide variety of different packs of synthetic and natural grains, (Zimmer, 2003; Zimmer et al., 2007) found exponents that were significantly $<1/6$. Recently, Andersen and Johansen (2010) found that Walton's (1987) model overpredicts the observed sonic log wave speeds in unconsolidated sands by more than a factor of 1.5, and they suggested empirically derived curves (e.g., Fam and Santamarina, 1997) may be best.

11.03.5.6 Implications of Pressure Dependence

The pressure sensitivity of moduli and wave speeds has two important implications. The first is that the application of nonhydrostatic or deviatoric stress states affects the rock's elastic anisotropy. The second is that the properties are also highly sensitive to a pore pressure. Both of these factors can strongly affect the seismic observations.

11.03.5.6.1 Stress-induced anisotropy (acoustoelastic effect)

All of the discussions to this point have assumed that the materials are subject to a uniform hydrostatic confining pressure P. In the Earth, however, a deviatoric state of stress is the general case. Such states of stress produce wave speed anisotropy in both granular (Sayers, 2002, 2007; Walton, 1987) and cracked (Horii and Nemat-Nasser, 1983; Sayers and Kachanov, 1995) rocks. In the former, the anisotropy results from the stiffening of the contacts in the direction parallel to the greater compression. In the latter, the anisotropy is primarily a consequence of the preferred closure of those cracks whose plane is aligned normal to the greatest principle compression. Nur and Simmons (1969) may have been the first to observe these phenomena although the stress concentrations in their circular sample subject to a uniaxial force were not uniform. Since then the effect has been documented by many workers under more controlled stress state conditions (e.g., Babuska and Pros, 1984; Becker et al., 2007; Bonner, 1974; Gurevich et al., 2011; Johnson and Rasolofosaon, 1996; Stanek et al., 2013) in cracked rock, (Nur, 1971) in granular rock (e.g., Dillen et al., 1999; Khidas and Jia, 2010; Prioul et al., 2004; Rai and Hanson, 1988; Roesler, 1979) and numerically (Gallop, 2013; Hu et al., 2010). At the field scale, Schijns et al. (2012) were able to explain the observed seismic anisotropy in a 1300 m thick section of the Outokumpu biotite gneiss on the basis of oriented microcracks and mineral CPO (Elbra et al., 2011; Kern et al., 2009; Lassila et al., 2010; Wenk et al., 2012).

The basis of this effect again derives from the progressive closure of the crack-like porosity with compression. Consider an unstressed rock mass containing numerous narrow cracks, the orientations of which are randomly but uniformly distributed (**Figure 22(a)**). Application of an appropriately large uniaxial stress σ_{xx} to the mass closes those cracks whose planes are perpendicular or nearly so to the direction of σ_{xx} (Berg, 1965; Walsh, 1965) while leaving the cracks whose planes are

aligned to the stress open as shown in Figure 22(b). This transforms the initially isotropic rock into a transversely isotropic one (Fuck and Tsvankin, 2009) with directionally dependent relative wave speeds as shown in Figure 22(c).

Stress-induced anisotropy is important in many geophysical problems from complicated stress conditions near a wellbore (Schmitt et al., 1989; Winkler, 1996), in mines (Holmes et al., 2000a; Holmes et al., 2000b), over petroleum reservoirs (Wuestefeld et al., 2011), and in stress changes related to seismicity. (Crampin, 1994; Crampin and Peacock, 2008) in particular had championed the interpretation of shear wave splitting to infer stress states for reservoir monitoring and earthquake forecasting.

11.03.5.6.2 Influence of pore pressure

Although we are not yet considering the effects on the dynamic seismic properties of fluids in the pore space of the rocks, the fluid pore pressure P_p has an important effect on the static elastic properties of the frame that must be taken account of in most cases. Numerous laboratory experiments have shown that the frame properties of rocks depend not on the confining pressure P per se but on the effective pressure P_e:

$$P_e = P - \xi P_p \qquad [78]$$

where ξ the pore pressure coefficient. More generally, for a set of total or confining stresses σ_{ij}, the effective stresses σ_{eij} eqn [78] can be written as

$$\sigma_{eij} = \sigma_{ij} - \delta_{ij}\xi P_p \qquad [79]$$

where δ_{ij} is the Kronecker delta; the pore fluid does not influence the shear stresses.

What this means is that the frame moduli K_d and μ_d and consequently the wave speeds V_P and V_S are generally functions of P_e. A corollary to this is that the dry moduli do not change (to the first approximation) if P_e remains constant. The measurements of He (2006) on the Cadotte sandstone illustrate these effects well (Figure 23). In his tests, he first measured V_P and V_S through a water-saturated sample of the Cadotte sandstone at confining pressures to 60 MPa while maintaining the pore pressure $P_p = 0$. This yields the highly nonlinear curve of velocities versus confining pressure as was already encountered for a similar dry sample in Figure 19. Once this was accomplished, He repeated the measurements in suites where the differential pressure was held constant first at 15 MPa and then at 30 MPa. In both of these, the wave speeds remain constant, meaning that $\xi = 1$ over this pressure range to within experimental error.

The concept of effective pressure can be confusing in part because a worker must take care to consider which definition of effective pressure one truly requires for the phenomena at hand. More directly, this usually means one must have the appropriate value of ξ for the physical process being studied (Berryman, 1992; Berryman, 1993). The effective pressure/stress concept was first developed by Terzaghi with $\xi = 1$. However, the *apparent* ξ will deviate from unity depending on the rock structure and potential of chemical interaction (e.g., swelling clays) (Bernabe, 1987; Zoback and Byerlee, 1975), on the rates at which pore pressures can recharge during rapid deformation and dilatancy leading to failure (Brace and Martin, 1968; Schmitt and Zoback, 1992), and on the linear strains and volumetric deformation of the material itself (Biot and Willis, 1957; Nur and Byerlee, 1971). With regard to the last item, the Biot–Willis effective stress coefficient *for poroelastic strain* is

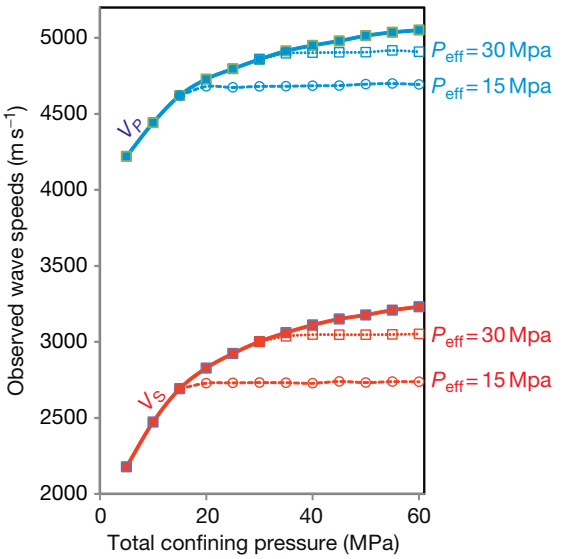


Figure 23 Observed ultrasonic velocities in a water-saturated sample of the Cadotte sandstone (see Figure 19) versus the total confining pressure. The solid lines are the velocities for V_P (blue) and V_S (red) for the sample with a small and constant pore pressure. Dashed and dotted lines and open circles and squares indicate the observations when the differential pressure $P - P_P$ is held constant at either 15 or 30 MPa, respectively. Data for sample SB007 of He (2006).

$$\xi = 1 - \frac{K_d}{K_m} \qquad [80]$$

Unfortunately, the use of eqn [80] in determining the appropriate effective pressure or stress in determining V_P or V_S has propagated through the rock physics literature. This usage is incorrect, as emphasized by Mavko and Vanorio (2010), and can lead to significant error particularly in compressible materials. For purposes of finding moduli or wave speeds, it is more proper to assume $\xi \approx 1$ in the absence of additional information on the material behavior.

The influence of effective pressure on seismic responses is of particular interest in time-lapse seismology. Pore pressures P_p in reservoirs will vary as fluids are injected or produced and this results in a change in the effective stress and, hence, V_P and V_S and the consequent seismic reflectivity. Proper interpretation of time-lapse observations cannot easily ignore the changes in the effective pressure or stress (Herwanger and Horne, 2005; Sayers, 2004).

11.03.6 Seismic Waves in Fluid-Saturated Rocks

Most commonly, a worker requires a prediction of the seismic wave speeds under the *in situ* conditions in the Earth where the rock is subject to various stresses, pore pressures, and fluid saturations. After providing information on the behavior of

the rock's various components of the mineral solid, the fluid, and the rock's frame in the preceding sections, we are now ready to review the approaches to understanding the seismic properties of fluid-saturated rocks. This is often referred to as a fluid substitution analysis. To summarize, the building blocks needed to determine the fluid-saturated bulk K_{sat} and shear μ_{sat} moduli that are necessary to calculate V_P and V_S with eqns [1] and [2] are as follows:

- The elastic moduli K_m and μ_m and density ρ_m of the solid mineral constituents.
- The saturating fluid's adiabatic bulk modulus K_f and its density ρ_f. If velocity dispersion is included, then one must also include the viscosity ζ.
- The rock frame's elastic dry or drained elastic moduli K_d and μ_d and its porosity ϕ. If frequency is considered, then one must also include additional factors such as the tortuosity τ, the permeability κ, and some measure of the dimensions of the pore.

Of course, one must remember that most of these depend on stress, pressure, and temperature as discussed in the previous sections and the values need to be appropriately chosen for the case at hand. After the intensive earlier discussions of the various mineral, fluid, and frame properties, the fluid-saturated expressions to come may seem anticlimactic. Often, finding the appropriate values of the inputs may take the most effort.

In this section, we review a number of the more important expressions that allow for the calculation of the saturated rock elastic moduli. This begins with Gassmann's widely employed formula and progresses through brief surveys of the local and global flow models that account for frequency effects.

11.03.6.1 Gassmann's Equation

Gassmann (1951) constructed expressions for the moduli of a fluid-saturated porous material. These expressions are relatively simple, and because of this, they are used almost exclusively in practice. His development is essentially that for an undrained poroelastic solid (Berryman, 1999; Rice and Cleary, 1976; Smith et al., 2003; Wang, 2000) and as such is strictly valid for static deformations. In developing his equations, he assumed the minerals and fluids interacted only mechanically, that the material was fully saturated and isotropic, that it was monomineralic, and, importantly, that the shear modulus is not influenced by the pore fluid. This last assumption leads to the expression for the saturated shear modulus μ_{sat}

$$\mu_{sat} = \mu_d \qquad [81]$$

and the saturated bulk modulus K_{sat}

$$K_{sat} = K_d + \frac{\xi^2}{\frac{\phi}{K_f} + \frac{\xi - \phi}{K_m}} \qquad [82]$$

where ξ is the poroelastic pore pressure coefficient of eqn [80]. Gassmann looked at the problem more as a superposition of interrelated volumetric strains of the frame, the minerals, and the fluid. The same result can be found independently as the undrained modulus from poroelasticity. Gassmann's formulas are written in many different ways, but one advantage of eqn [82] is that it highlights that the fluid effects are essentially a correction to K_d (Han and Batzle, 2004).

For the sake of convenience it is also worth rewriting an inverted form of Gassmann's equation [82]:

$$K_d = \frac{\left(\frac{\phi K_m}{K_f} + 1 - \phi\right) K_{sat} - K_m}{\frac{\phi K_m}{K_f} - 1 - \phi + \frac{K_{sat}}{K_m}} \qquad [83]$$

which is particularly useful in estimating K_d from sonic log information (Carcione et al., 2006).

Physically, the application of Gassmann's equations [81] and [82] assumes that during the passage of a wave over a small (relative to the wavelength) but representative volume V^R of the saturated rock, the frequency is sufficiently low that the pore pressure remains uniform throughout V^R. With no variations in pressure, there can be no fluid exchange between V^R and its neighboring representative volumes.

There are a number of salient points arising from Gassmann's equations including the following:

i. $K_{sat} \geq K_d$. That is, a fluid-saturated rock is always less compressible than the dry frame by itself.
ii. $1 \geq \xi \geq 0$. The upper bound occurs as K_d vanishes for increasingly compressible materials such as unconsolidated sands. It is interesting to note by examination that as $\xi \to 1$, K_{sat} approaches the Voigt bound equation [58] for the mixture of fluid and solid. The lower bound is approached as $K_d \to K_m$ where the rock is very stiff. Consequently, the influence of fluids diminishes as the frame stiffens.
iii. $K_m \gg K_f$. As **Tables 1** and **2** show, typical mineral moduli are mostly >30 GPa. In contrast, most typical liquid moduli are $\sim$1–3 GPa and gas moduli are $\sim$0.1–100 MPa (**Table 3**). As such, for real materials, K_{sat} is substantially more sensitive to K_f than to K_m. A corollary to this is that a compressible fluid (i.e., gas) has less effect than does a stiffer fluid (liquid).
iv. V_S will decrease upon saturation. The decrease, although it may only be small, occurs because, by definition, the shear modulus does not change but the bulk density ρ as given in eqns [55] and [56] must increase.
v. V_P will usually, but not necessarily, increase upon saturation. This increase is primarily due to the larger value of K_{sat}. However, one must always keep in mind that ρ_{sat} also rises, and should ρ_f be sufficiently large, it is quite possible that V_P will decrease upon saturation. This happens, for example, with pure CO_2 (e.g., see Njiekak et al., 2013).

It is worthwhile investigating the implications of Gassmann's equations. **Figure 24(a)** shows a suite of dry elastic moduli that might represent the values for a number of limestones with a range of porosities $0 \leq \phi \leq 0.4$. Both K_d and μ_d decrease substantially with porosity as expected. K_{sat} is calculated using eqn [82]; it is always greater than K_d but the two continue to diverge from one another with porosity and K_{sat} is nearly twice K_d once a porosity $\phi = 0.35$ is reached. The dry and saturated shear moduli are exactly the same according to Gassmann's assumptions in eqn [81].

The seismic wave speeds may subsequently be calculated using eqns [1] and [2] (**Figure 24(b)**). V_S decreases a small

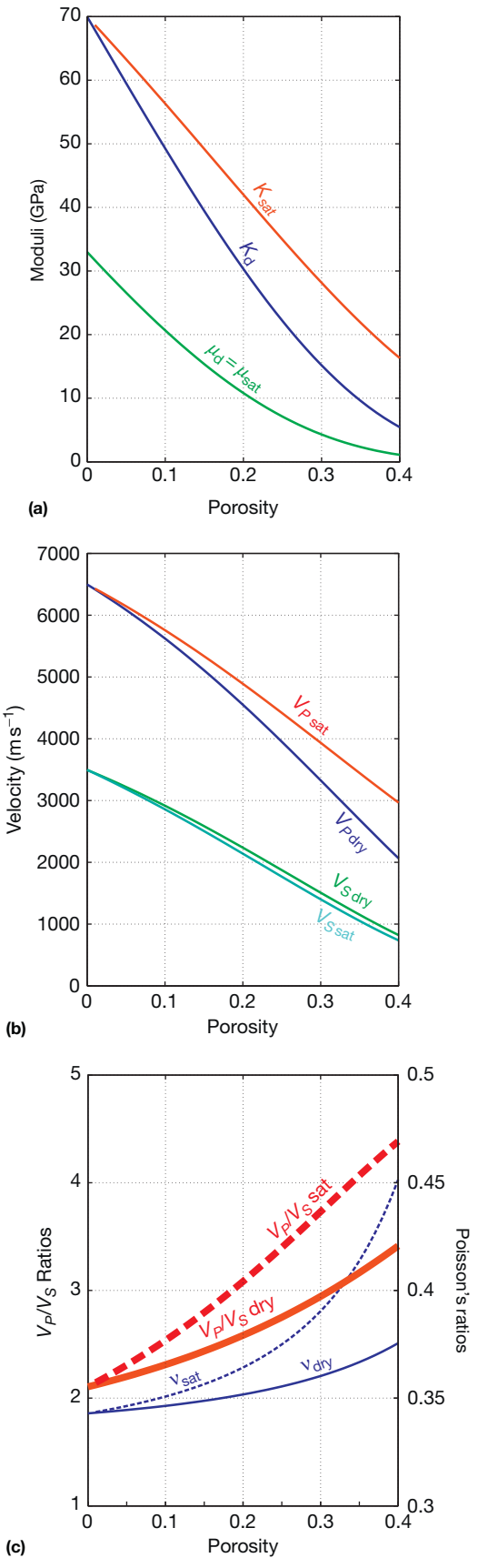


Figure 24 Illustration of the application of Gassmann's equations to a series of hypothetical rocks of increasing porosity. (a) Comparison of the dry frame modulus bulk K_d (blue) to the saturated bulk modulus K_{sat}. Note that by definition, the dry μ_d and saturated μ_{sat} shear moduli are equal. (b) Corresponding compressional and shear waves speeds for the dry and saturated cases. (c) Corresponding V_P/V_S and Poisson's ratios for the dry and saturated cases.

amount upon substitution because of the increase in the bulk density as already noted. For this case, however, V_P increases relative to that for the dry frame; the increase in the bulk modulus has managed to overwhelm the increase in the density. At the greater porosities (or equivalently smaller K_d), the saturating water has a greater influence on V_{Psat}. These differences would be further amplified for the acoustic impedances. It is these types of differences that hold the key to the time-lapse seismic surveying.

Figure 24(c) shows the two parameters that are often derived from the wave speeds, the direct V_P/V_S ratio and Poisson's ratio equation [9]. These can serve as proxy indicators of the saturation state of the rock in part because of the relative impacts of saturation on V_{Psat} and V_{Ssat}. For example, the V_P/V_S ratio for water saturation is substantially greater than that for the dry (or nearly equivalently gas-filled) conditions. This difference is even greater for Poisson's ratio ν with a larger Poisson's ratio, indicating liquid saturation.

11.03.6.2 Frequency-Dependent Models

Gassmann's equations remain popular, but as noted, they strictly apply only under static conditions. However, in reality, researchers must work over a large range of frequencies. In the broadest sense, 'seismic' investigations center on frequencies of 1 MHz, 10 kHz, and 100 Hz for ultrasonic laboratory measurements, borehole sonic log readings, and applied seismic investigations, respectively.

In the laboratory, Gassmann's equations rarely reproduce the observed wave speeds particularly at low confining pressures. An example of this, again taken from measurements on the Cadotte sandstone (He, 2006), is given in **Figure 25** that shows the observed dry and water-saturated measurements as a function of the effective confining pressure. The observed V_{Psat} are greater than V_{Pdry} as expected but they exceed those values predicted using Gassmann's equations significantly. Similar anomalies are seen for V_{Ssat} that again is much greater than V_{Sdry}; this is in opposition to Gassmann's predicted values that are slightly smaller than V_{Sdry} because of the increased density of the bulk saturated material. It is important to note that as the crack-like porosity closes, the observed and predicted values begin to approach one another. Indeed, even at the pressure of only 60 MPa, V_{Ssat} has nearly intersected with V_{Sdry}. Clearly, there is a large discrepancy between the observed and predicted saturated wave speeds particularly at the low confining pressures where the cracks remain most open (see also Mavko and Jizba (1991) for similar evidence).

An extreme case of wave speed dispersion is seen from recent V_P and V_S measurements through a highly porous, CO_2-saturated medium of sintered alumina (Yam, 2011; Yam and Schmitt, 2011). **Figure 26** shows the variations of the observed ultrasonic (1 MHz) V_{Psat} and V_{Ssat} with pore pressure P_p for a suite of measurements carried out at constant differential effective stress $P_{eff} = P - P_p$ of 15 MPa. As noted in the

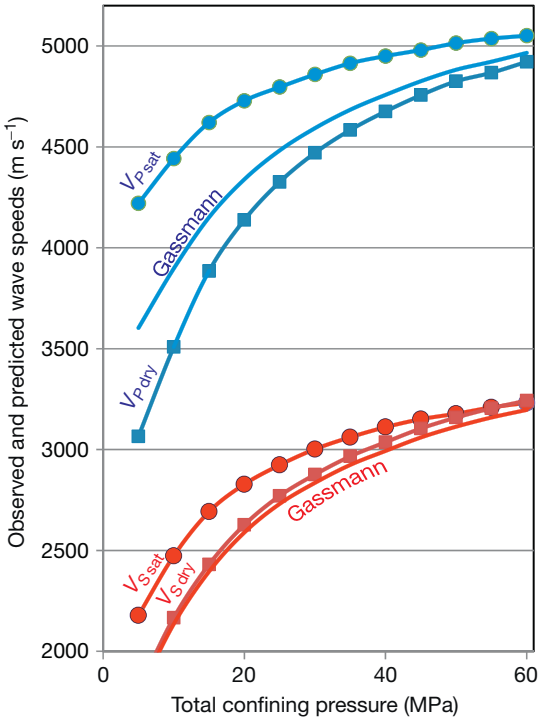


Figure 25 Observed V_P (blue) and V_S (red) wave speeds in Cadotte sandstone SB007 for dry (filled squares) and water-saturated (filled circles) situations. Theoretical Gassmann's wave speeds calculated using the dry moduli are also shown as lines.

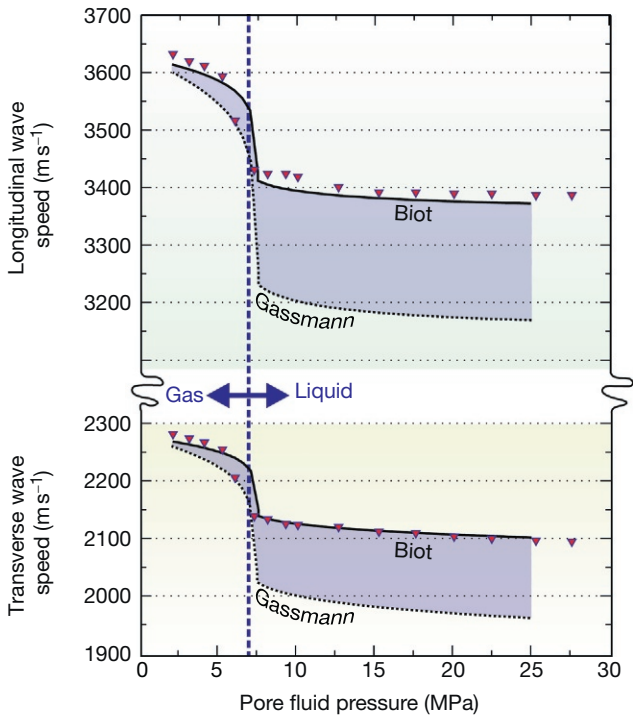


Figure 26 Comparison of observed ultrasonic wave speeds in a CO_2-saturated porous sintered alumina held at a constant differential pressure of $P_{eff} = 15$ MPa and temperature $T = 28$ °C. The pore pressure is varied in order to effect the gas–liquid phase transition that occurs at 6.144 MPa. Observed values of V_{Psat} and V_{Ssat} are indicated by red upside down triangles. Calculated values of the wave speeds according to the Biot formulations at 100 and 1 MHz are shown as dashed and solid black lines, respectively. For this case, the fluid substitution calculations using Gassmann's zero frequency equations cannot be distinguished from the 100 Hz Biot calculations. The violet-filled zones indicate the range of wave speed dispersion that exists between 100 Hz (Gassmann) and 1 MHz (Biot). Data and calculations from Yam (2011).

preceding text, carrying out the measurements at constant differential pressure should minimize any pressure-dependent effects of the rock frame of wave speeds and most of the variation will be due to changes in the fluid properties. Figure 26 also shows the corresponding wave speeds predicted by Gassmann's equations from fluid substitution calculations that employ knowledge of the fluid, solid, and dry frame properties. Gassmann's predictions are always less than the observations. This difference is relatively small when CO_2 gas saturates the pores but it grows to over 6% when the CO_2 is in the liquid state.

In the case shown, the differences between the observed and the predicted wave speeds at low frequencies arise primarily because the latter does not account for differential motions between the solid and fluid as the wave passes. We briefly survey the two main models in the succeeding text that do admit fluid motions. These are usually referred to as the global and local flow models.

11.03.6.2.1 Global flow (biot) model

Building on his work in the consolidation of porous media, Biot (1956a,b) constructed a seminal model to account for frequency-dependent wave propagation through a fluid-saturated porous and permeable medium. A full description of the model is beyond the needs here but readers can find more information in numerous contributions (Bourbié et al., 1987; Johnson, 1984; Smeulders, 2005). A most interesting outcome of this theory is that a third 'slow' longitudinal wave mode P_2 exists in such porous media because the fluid and solid are allowed to move independently of one another. This out-of-phase motion between the two admits a new degree of freedom to the equations of which the slow wave speed is the consequence.

This wave mode is difficult to observe in nature because the differential fluid–solid motions result in large viscous losses. Plona (1980) first observed this wave in acoustic refraction experiments through highly porous sintered glass beads. Figure 27 shows the results from similar recently conducted tests (Bouzidi and Schmitt, 2009) that highlight well the existence of all three modes. It is interesting to note the loss of the fast P_1 wave past its critical angle. The P_2 mode, which travels substantially more slowly than the surrounding water, has no critical angle and exists at all angles.

At this point, it is important to return to the CO_2-saturated measurements of Figure 26 where the Biot model has been used to predict the observed V_{Psat} and V_{Ssat}. The theory and observations match well as the observed and calculated values of V_{Ssat} agree to within experimental uncertainty. Those for V_{Psat} are only 20 m s^{-1} different. This agreement strongly suggests that the Biot model and its low-frequency limit given by Gassmann's equations adequately predict the behavior with frequency in highly porous and permeable materials. Although not shown here, the Biot formulas also allowed attenuation to be modeled satisfactorily (Yam, 2011).

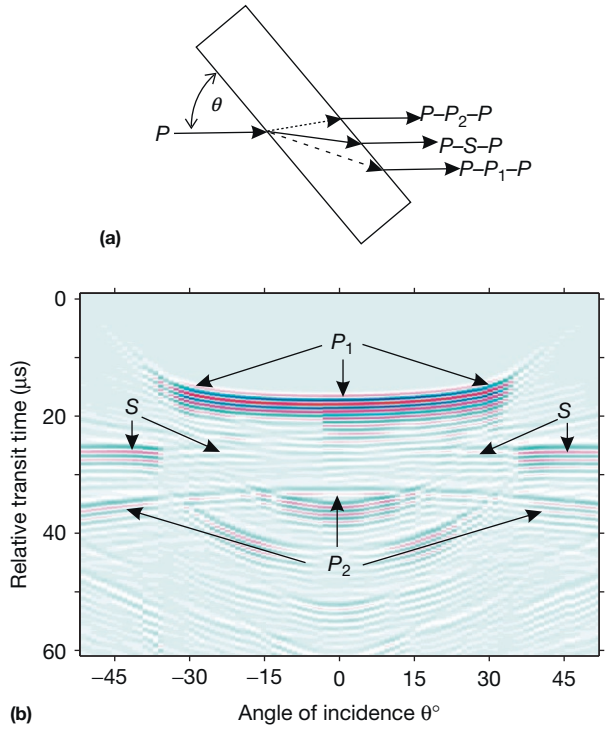


Figure 27 Experimental arrangement to show the existence of the slow wave. (a) Schematic of the experiment. An ultrasonic pulse insonifies a saturated porous plate immersed in a water tank with incoming longitudinal wave *P*. At the first boundary of the sample, the *P* wave converts to the 'fast' P_1, the *S*, and the 'slow' P_2 modes that propagate through the sample. These modes then are again converted back to separate *P* waves at the second surface and these waterborne arrivals are detected by an ultrasonic transducer. (b) Observed ultrasonic waveforms as a function of the incidence angle θ of the incoming waterborne *P* wave.

However, there are two problems with the use of the Biot formulas. First, properly carrying out the analysis requires knowledge of at least 14 different physical properties and characteristics of the porous medium. This is regrettably impractical in many situations. Second, the model does not account for all differential fluid motions within the porous rock and fails in predicting the wave speeds and attenuation through rocks containing compressible crack-like pores. This situation is discussed next.

11.03.6.2.2 Local flow (squirt) models

Although the concept was initially suggested by Mavko and Nur (1975), O'Connell and Budiansky (1977) carried out a theoretical analysis of a cracked solid containing cracks filled with a viscoelastic fluid. They noted that as long as the cracks are connected, then application of a stress to the porous medium allows the fluids to flow locally from those cracks preferentially oriented such that they compress (see **Figure 22**) into those cracks oriented more parallel to the stress. That is, the pressurized fluid in the compressed cracks 'squirts' into the uncompressed cracks. These fluid motions will preferentially occur at frequencies controlled by the crack density, effective permeability, and crack dimensions. At sufficiently high frequencies, the fluids cannot move fast enough and essentially become locked within the pores causing the medium to become stiffer. Similar phenomena can exist if the material contains both compressible cracks and stiffer equant pores.

O'Connell and Budiansky (1977) described a number of frequency domains that depend on the relationships between crack dimensions, frequency, and viscosity. Jackson (Lu and Jackson, 2006) had provided a particularly useful summary of their limiting cases that from high to low frequency are the *glued* (i.e., shear stresses do not have time to relax), the *saturated isolated* (i.e., shear stresses in the fluids relax but there is insufficient time for fluid exchange between pores), the *saturated isobaric* (i.e., adjacent pores can exchange fluids but there is no global fluid movement; this is similar to the undrained case of Gassmann's equations), and, finally, the *drained* case (i.e., bulk fluids can move in or out of the material to equilibrate the pore pressure and the moduli act as if the sample were dry).

These shifts between fluid-flow regimes impact the moduli and attenuation in different ways as illustrated in **Figure 28**. A large shift in the bulk modulus K occurs for the transition from the drained to undrained conditions. At low frequencies, there is no dispersion in the shear moduli μ across this same transition. Once viscous forces come into play, however, the shear moduli stiffen in the transitions from the saturated isobaric to the saturated isolated and from the saturated isolated to the glued regimes. A peak in the attenuation is expected at each of these jumps in the moduli with higher frequency.

These developments have been useful conceptually but finding appropriate expressions to account for these local flow effects has remained challenging. One issue is that the actual dimensions of the cracks play a significant role but appropriately assessing crack sizes is not easily accomplished. Mavko and Jizba (1991) attempted to indirectly account for the local flow effects by measuring the 'crack-free' modulus at high confining pressures in order to provide an estimate of the 'wet frame modulus.' These moduli are then used in Gassmann's equations to provide for a prediction of the high-frequency moduli. Work continues on this issue with numerous recent contributions (Adelinet et al., 2011; Dvorkin et al., 1995; LeRavalec and Gueguen, 1996; Schubnel and Gueguen, 2003).

11.03.7 Empirical Relations and Data Compilations

The sections in the preceding text have attempted to give some sense of the problems associated with predicting the seismic properties of rocks. Often, however, there is far from sufficient information in order to justifiably carry out such predictions, and workers rely instead on the observations of others. There are numerous drawbacks to this. For example, from the perspective of geology, the name of a rock type may reveal a great deal about how a rock was formed, what minerals might be in it, and what its texture is. However, the geologic name by itself may have little meaning from the perspective of the rock's physical properties except in the broadest of senses. Despite this, sometimes this is the only information that we have available.

Mavko et al. (2003) had already extensively reviewed a number of existing empirical relationships for a wide variety of rock types; as such we do not need to reproduce all of these

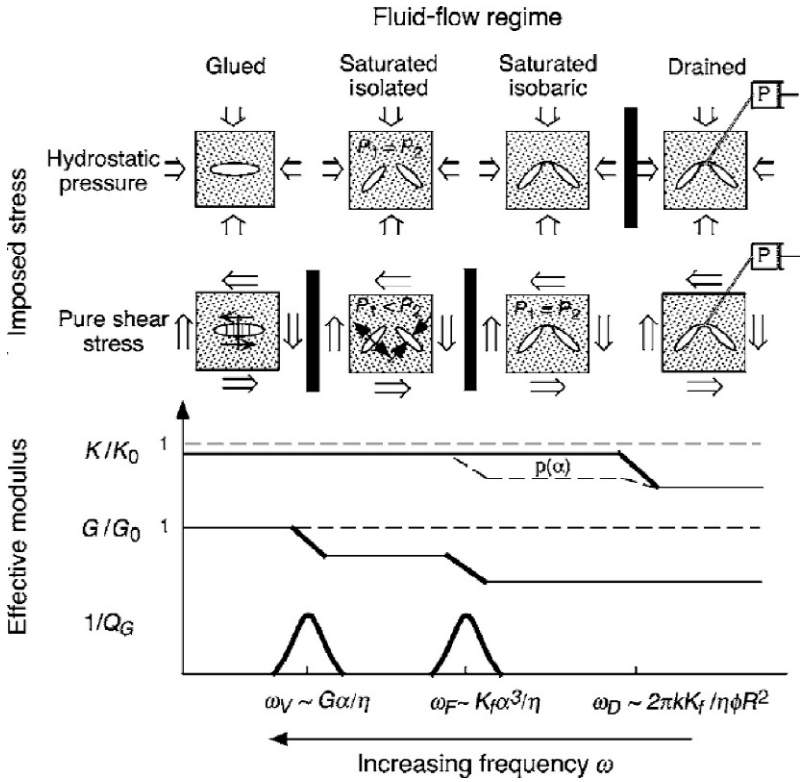


Figure 28 Description of the fluid-flow regimes expected in a cracked and fluid-saturated solid during passage of a harmonic wave according to O'Connell and Budiansky (1977) as explained by Lu and Jackson (2006). Figure 10 from Lu and Jackson (2006) with permission granted according the Society of Exploration Geophysicists fair use policy.

here, but we provide a few that have been popular in the literature. Dvorkin (2008) also provided a listing of many of the simpler linear fits for V_P and V_S.

Eberhart-Phillips et al. (1989) fitted various functions to the large ultrasonic data set of saturated shaley sandstones obtained by Han (1986) and arrived at the forms

$$V_P = 5.77 - 6.94\phi - 1.73\sqrt{C} + 0.446(P_{\text{eff}} - \exp(-16.7P_{\text{eff}}))$$
$$V_S = 3.70 - 4.94\phi - 1.57\sqrt{C} + 0.361(P_{\text{eff}} - \exp(-16.7P_{\text{eff}})) \quad [84]$$

where V_P and V_S are in km s^{-1}, C is the clay volume fraction, and P_{eff} is the differential pressure in kilobars (1 kbar = 100 MPa).

Castagna et al. (1985) derived what they called a 'mudrock' line from sonic logs allowing V_S to be predicted from V_P for siliclastics using

$$V_S = 0.862V_P - 1.1172 \quad [85]$$

where both velocities are in km s^{-1}. This simple linear fit has surprising predictive power. Jorstad et al. (1999) later carried out their own regression for their area of study and produced a remarkably similar relation:

$$V_S = 0.8966V_P - 1.1665 \quad [86]$$

Following the earlier work of Pickett (1963), Castagna et al. (1993) also developed a set of quadratic regression polynomials relating V_S to V_P for the pure lithologies of sandstone, limestone, dolomite, and shale saturated fully with brine of the form

$$V_{Si} = a_{i2}V_P^2 + a_{i1}V_P + a_{i0} \quad [87]$$

with the resulting empirical coefficients a_{i0}, a_{i1}, and a_{i2} provided in Table 4. Note that except for limestone, a linear fit suffices to describe the V_S–V_P relationships.

Table 4 Quadratic regression coefficients for pure rock types of Castagna et al. (1993)

Rock type	a_{i2}	a_{i1}	a_{i0}	R^2
Sandstone	0	0.80416	−0.8558	0.98352
Limestone	−0.05508	1.01677	−1.3049	0.99096
Dolomite	0	0.58324	−0.07775	0.87444
Shale	0	0.76969	−0.86735	0.97939

Greenberg and Castagna (1992, 1993) developed an iterative method that allows mixing of pure monomineralic 'lithologies' of sandstone (quartz), limestone (calcite), dolomite (dolomite), and shale (illite) lithologies together to predict V_S for variable fluid saturations. Their method is somewhat analogous to Hill's mean value of the Voigt–Reuss bounds as described in **Section 11.03.5.2**. Essentially, the strategy is to find the 'best' predicted value of the compressional wave velocity V_{P1C} of the rock under study when it is brine-saturated. Once this is found, the equivalent brine-saturated shear wave

velocity is calculated using the empirical regressions of Equation [87] just described. With knowledge of the brine-saturated density, one can easily calculate the rock's shear modulus μ_d that in turn can be used to calculate the proper shear wave velocity for the saturation at the proper saturation state using the rock's density ρ_{sat}.

Hence, the basis of the method is to make a proper estimate of V_{P1C}. To carry this out, one must have appropriate knowledge of the rock's porosity ϕ, its *in situ* saturations S_w, S_o, and S_g (see eqns [56] and [57]), and its observed *in situ* compressional velocity V_{Psat}. Further, the solid volume fractions X_i of the L different 'lithologies'

$$\sum_{i=0}^{L} X_i = 1 \qquad [88]$$

forming the rock's frame must also be found.

In the Greenberg and Castagna method, they first guess a value of the compressional wave speed V_{P1C} for the rock fully saturated with brine (i.e., $S_w = 1$). With the guess V_{P1C}, one can directly calculate the expected values V_{Si} for each of the lithologies using eqn [87] and **Table 4**. Using these V_{Si}, the shear wave velocity for the brine-saturated rock V_{S1C} is calculated in a manner analogous to the Voigt–Reuss–Hill calculations:

$$V_{S1C} = \frac{1}{2}\left\{\left[\sum_{i=0}^{L} X_i V_{Si}\right] + \left[\sum_{i=0}^{L} \frac{X_i}{V_{Si}}\right]^{-1}\right\} \qquad [89]$$

The density of the brine-saturated rock ρ_{1C} is calculated using an adaptation of eqn [55]:

$$\rho_{1C} = \phi\rho_w + (1-\phi)\sum_{i=1}^{L} X_i \rho_{mi} \qquad [90]$$

where the brine density must be determined for the appropriate conditions as discussed in **Section 11.03.4.3.2.1**. Note that the Greenberg and Castagna (1992) lithologies are sandstone, limestone, dolomite, and shale. The mineral densities ρ_{mi} that they employed for these lithologies are, respectively, those for quartz, calcite, dolomite, and illite as could be obtained from **Table 2**. The *in situ* saturated density ρ_{sat} can be either directly observed or calculated with an additional adaptation of eqn [56]:

$$\rho_{sat} = \phi\left(\sum_{j=1}^{n} S_j \rho_j^f\right) + (1-\phi)\sum_{i=1}^{L} X_i \rho_{mi} \qquad [91]$$

With the values of V_{S1C} and ρ_{1C}, the drained shear modulus μ_d using eqn [2]

$$\mu_d = V_{S1C}^2 \rho_{1C} \qquad [92]$$

may be calculated. The saturated bulk modulus K_{sat} of the rock in situ with its proper saturation is then from eqn [1]:

$$K_{sat} = V_{Psat}^2 \rho_{sat} - \frac{4}{3}\mu_d \qquad [93]$$

The drained frame modulus K_d may then be determined using the inverse Gassmann's equation [83] with knowledge of the appropriate K_f for the pore fluids and K_m for the mineral solids discussed in earlier sections. Now, the calculations are run somewhat backward in order to provide an updated value for the fully brine-saturated ($S_w = 1$) compressional wave speed V'_{P1C} using eqn [1]:

$$V'_{P1C} = \left\{\left[\left(K_d + \frac{\xi^2}{\frac{\phi}{K_w} + \frac{\xi-\phi}{K_m}}\right) + \frac{4}{3}\mu_d\right]\rho_{1C}^{-1}\right\}^{1/2} \qquad [94]$$

with the bulk modulus of the brine K_w taking the place of K_f in Gassmann's equation [82]. V'_{P1C} can then be used as the new V_{P1C} and the calculation repeated. This process can iterate until V'_{P1C} and V_{P1C} converge within acceptable limits. Once the convergence is complete, the current value of μ_d is accepted and the final estimate of the shear wave velocity V_{Ssat} is simply calculated using eqn [2]:

$$V_{Ssat} = \sqrt{\frac{\mu_d}{\rho_{sat}}} \qquad [95]$$

Raymer et al. (1980) also using well log information developed a piecewise relationship to predict velocity from knowledge of porosity and the intrinsic compressional wave speeds V_{Pm} and V_f for the constituent mineral and solid, respectively. This formula is

$$\begin{aligned} V_P &= (1-\phi)^2 V_{Pm} + \phi V_f, \ \phi < 0.37 \\ \frac{1}{\rho V_P^2} &= \frac{\phi}{\rho_f V_f^2} + \frac{1-\phi}{\rho_m V_{Pm}^2}, \ \phi > 0.47 \\ \frac{1}{V_P} &= \frac{0.47-\phi}{0.10}\frac{1}{V_{37}} + \frac{\phi-0.37}{0.10}\frac{1}{V_{47}}, \ 0.37 < \phi < 0.47 \end{aligned} \qquad [96]$$

where V_{37} and V_{47} are the velocities of the saturated rock for $\phi = 37\%$ and $\phi = 47\%$, respectively. Dvorkin (2008) more recently came up with a similar relation to predict shear wave velocities:

$$V_S = (1-\phi)^2 V_{Sm}\left[\frac{(1-\phi)\rho_m}{(1-\phi)\rho_m + \phi\rho_f}\right] \qquad [97]$$

given knowledge of the intrinsic mineral shear wave velocity V_{Sm} and the fluid and intrinsic mineral densities, ρ_f and ρ_m, respectively.

Hamilton (1979) derived a number of expressions for velocities in marine sediments (primarily silt–clays, turbidites, and mudstone shales); one of these relates V_P (in km s^{-1}) to depth D (in km) by

$$V_P = 1.511 + 1.304D - 0.741D^2 + 0.257D^3 \qquad [98]$$

which is accurate to about 800 m depth.

The general empirical relationship between density and seismic velocity fit by Gardner et al. (Gardner et al., 1974; Sayers and den Boer, 2011) is still widely applied in seismic studies particularly if only one of the other logs is available from a borehole

$$\rho \approx 1.741 V_P^{0.25} \qquad [99]$$

with V_P in km s^{-1} and ρ in g cm^{-3}.

More recently, Christensen and Stanley (2003) carried out a fit of V_P and V_S versus ρ from a compilation of laboratory measurements on a variety of igneous, metamorphic, and sedimentary rocks yielding

$$V_P = 2.8\rho - 1.5752 \qquad [100]$$

and

$$V_S = 1.4\rho - 0.4019 \qquad [101]$$

where again the velocities are in km s^{-1} and ρ in g cm^{-3}.

Examination of these regressions shows that, for the most part, they do not contain any physical principles. Hence, the reader must take care when applying these formulas and know their limitations.

Compilations of observed values of velocities and densities are also important sources of information. The densities of rocks and minerals are covered extensively by Wohlenberg (1982), while in the same series, Gebrande (1982) and Kern (1982) provided exhaustive listings of wave speeds as functions of pressure for various rock types. Christensen (1982, 1989) had provided the results of thousands of ultrasonic velocity laboratory measurements that have been used to create sophisticated models of the continental crust (Christensen and Mooney, 1995). Ji et al.'s (2002) compilation is perhaps the most complete in that it contains the results of both modeling and measurements for a wide variety of different rock types. Most recently, Gercek (2007) had compiled values of Poisson's ratio for a number of different Earth materials.

11.03.8 The Road Ahead

Studies related to 'seismic properties' continue to rapidly expand, and in reality, each of the topics touched on in the preceding text could be easily be turned into textbooks on their own. In writing such a review, some important issues were overlooked. The listing in the succeeding text contains important directions of investigation that I believe will be of increasing importance in the coming decade. These include the following:

i. Seismic anisotropy was touched on only in the context of mineral properties, but in general, rock is an anisotropic medium. Since the 1980s, geophysicists have dominated the literature on elastic anisotropy. As noted earlier, we do not really have a good understanding of the anisotropy of 'shales.' This is in part because the physical properties of clay mineralogical constituents are still not well known. Characterizing the clay mineralogy within a given sample remains problematic. The clay minerals themselves are difficult to study as the 'crystals' are often no more than a few nanometers in dimension. The effects of cracks and fractures on seismic responses still remain elusive. Being able to deconvolve seismic observations to detect and characterize fractures remains a goal for the future.
ii. Although it has been hinted at in the discussions earlier in the text, we currently do not have a good handle on seismic velocity dispersion in the Earth. Progress in this field should come more rapidly with the development of a variety of forcing-function stress–strain techniques that can allow the complex elastic moduli to be found at seismic frequencies.
iii. We have seen earlier that most rocks display nonlinear stress–strain or stress–velocity behavior. Despite this, we still cling to attempting to analyze much of our observed seismic data under the linearly elastic assumption. Currently, only a handful of workers are studying the often unexpected effects of wave propagation in nonlinear materials. It is likely that this field can gain from the much larger community of workers in nonlinear optics.
iv. In the laboratory, we often work on decimeter-scale samples that appear uniform and homogeneous. We then attempt to extend this assumption of homogeneity to much larger scales over many tens of meters (about the seismic wavelengths). However, there are subtle dispersion effects introduced by the scale of the heterogeneity of a rock mass relative to the interrogating seismic waves. For example, we most usually work under the ray theory paradigm that assumes the wavelengths are small relative to the dimensions of the heterogeneity. This is likely only rarely true, and in many cases, the waves we observe may more actually be revealing an effective medium. Much work remains to be done on assessing to what degree effective medium concepts may need to be incorporated to our field observations.
v. Finally, the remarkable growth in computing power and in easily employed parallel multiphysics numerical modeling together with parallel technical developments in benchtop x-ray tomography systems with micrometer-scale resolving power is leading to the rapid growth of 'digital rock physics.' These techniques allow one to calculate a variety of physical properties from the 3-D microtomograms. Estimation of electrical conductivity and permeability has been most successful so far, but accurate predictions of the elastic moduli and wave speeds have not yet been adequately solved. This is in part because for the most part, workers have focused on continuum approaches to solving such problems. Now, however, one can build actual models mimicking the true architecture of the minerals and pore space and more thought needs to go on with regard to actual physical phenomena in porous media at that scale before the seismic properties can be adequately modeled.

Taken together, there still remains a great deal to learn with regard to seismic properties of rocks. The new computational, imaging, and experimental tools now at our disposal will allow for rapid progress over the next decade.

Acknowledgments

The help of the many graduate students, postdoctoral researchers, and technical staff in making many of the measurements described in the preceding text is greatly appreciated. DRS is supported by the Canada Research Chairs Program. The author thanks T. Smith and the Editor L. Slater for providing ideas on improving the manuscript.

References

Adams LH and Williamson ED (1923) On the compressibility of minerals and rocks at high pressures. *Journal of the Franklin Institute* 195: 475–529.

Adelinet M, Fortin J, and Gueguen Y (2011) Dispersion of elastic moduli in a porous-cracked rock: Theoretical predictions for squirt-flow. *Tectonophysics* 503: 173–181.

Alam MM, Fabricius IL, and Christensen HF (2012) Static and dynamic effective stress coefficient of chalk. *Geophysics* 77: L1–L11.

Alexandrov KS and Ryzhova TV (1961a) The elastic properties of rock forming minerals I: Pyroxenes and amphiboles. *Bulletin (izvestiya), Academy of Sciences, USSR. Geophysics Series*, vol. 9, pp. 871–875.

Alexandrov KS and Ryzhova TV (1961b) The elastic properties of rock forming minerals II: Layered silicates. *Bulletin (izvestiya), Academy of Sciences, USSR. Geophysics Series*, vol. 12, pp. 1799–1804.

Alexandrov KS and Ryzhova TV (1962) The elastic properties of rock forming minerals III: Feldspars. *Bulletin (izvestiya), Academy of Sciences, USSR. Geophysics Series*, vol. 2, pp. 186–189.

Alexandrov KS and Ryzhova TV (1964) The elastic properties of pyroxenes. *Soviet Physics – Crystallography* 8: 589–591.

Almqvist BSG, Burlini L, Mainprice D, and Hirt AM (2010) Elastic properties of anisotropic synthetic calcite-muscovite aggregates. *Journal of Geophysical Research, Solid Earth* 115: B08203.

Andersen CF and Johansen TA (2010) Test of rock physics models for prediction of seismic velocities in shallow unconsolidated sands: A well log data case. *Geophysical Prospecting* 58: 1083–1098.

Anderson OL (1963) A simplified method for calculating debye temperature from elastic constants. *Journal of Physics and Chemistry of Solids* 24: 909–917.

Angel RJ, Jackson JM, Reichmann HJ, and Speziale S (2009) Elasticity measurements on minerals: A review. *European Journal of Mineralogy* 21: 525–550.

Angus DA, Verdon JP, Fisher QJ, and Kendall JM (2009) Exploring trends in microcrack properties of sedimentary rocks: An audit of dry-core velocity-stress measurements. *Geophysics* 74: E193–E203.

Auld BA (1990) Acoustic Fields and Waves in Solids, 2nd edn., p. 432. New York: John Wiley and Sons.

Auzende AL, Pellenq RJM, Devouard B, Baronnet A, and Grauby O (2006) Atomistic calculations of structural and elastic properties of serpentine minerals: The case of lizardite. *Physics and Chemistry of Minerals* 33: 266–275.

Azeemuddin M, Scott TE, Zaman M, and Roegiers JC (2001) *Stress-Dependent Biot's Constant Through Dynamic Measurements on Ekofisk Chalk*. Paper presented at 38th Symposium on Rock Mechanics: DC Rocks 2001. New York: Balkema.

Babuska V (1972) Elasticity and anisotropy of dunite and bronzitite. *Journal of Geophysical Research* 77: 6955–6965.

Babuška V and Cara M (1991) *Seismic Anisotropy in the Earth*. Netherlands: Springer.

Babuska V and Pros Z (1984) Velocity anisotropy in granodiorite and quartzite due to the distribution of microcracks. *Geophysical Journal of the Royal Astronomical Society* 76: 113–119.

Bachrach R and Avseth P (2008) Rock physics modeling of unconsolidated sands: Accounting for nonuniform contacts and heterogeneous stress fields in the effective media approximation with applications to hydrocarbon exploration. *Geophysics* 73: E197–E209.

Bachrach R, Dvorkin J, and Nur AM (2000) Seismic velocities and poisson's ratio of shallow unconsolidated sands. *Geophysics* 65: 559–564.

Bailey E and Holloway JR (2000) Experimental determination of elastic properties of talc to 800 degrees C, 0.5 GPa; Calculations of the effect on hydrated peridotite, and implications for cold subduction zones. *Earth and Planetary Science Letters* 183: 487–498.

Bakhorji A (2009) *Laboratory Measurements of Static and Dynamic Elastic Properties in Carbonate*. PhD, University of Alberta.

Bass JD (1995) Elasticity of minerals, glasses, and melts. In: Ahrens TJ (ed.) *Mineral Physics and Crystallography: A Handbook of Physical Constants*, pp. 45–63. Washington, DC: American Geophysical Union.

Batzle M and Wang ZJ (1992) Seismic properties of pore fluids. *Geophysics* 57: 1396–1408.

Becker K, Shapiro SA, Stanchits S, Dresen G, and Vinciguerra S (2007) Stress induced elastic anisotropy of the Etnean basalt: Theoretical and laboratory examination. *Geophysical Research Letters* 34: L11307.

Behura J, Batzle M, Hofmann R, and Dorgan J (2007) Heavy oils: Their shear story. *Geophysics* 72: E175–E183.

Belogol'skii VA, Sekoyan SS, Samorukova LM, Stefanov SR, and Levtsov VI (1999) Pressure dependence of the sound velocity in distilled water. *Measurement Techniques* 42: 406–413.

Benedict M, Webb GB, and Rubin LC (1942) An empirical equation for thermodynamic properties of light hydrocarbons and their mixtures II. Mixtures of methane, ethane, propane, and *n*-butane. *Journal of Chemical Physics* 10: 747–758.

Benedict M, Webb GB, and Rubin LC (1951) An empirical equation for thermodynamic properties of light hydrocarbons and their mixtures – Constants for 12 hydrocarbons. *Chemical Engineering Progress* 47: 419–422.

Berg CA (1965) Deformation of fine cracks under high pressure and shear. *Journal of Geophysical Research* 70: 3447–3452.

Berge PA, Bonner BP, and Berryman JG (1995) Ultrasonic velocity porosity relationships for sandstone analogs made from fused glass-beads. *Geophysics* 60: 108–119.

Bernabe Y (1987) The effective pressure law for permeability during pore pressure and confining pressure cycling of several crystalline rocks. *Journal of Geophysical Research-Solid Earth and Planets* 92: 649–657.

Berryman JG (1992) Effective stress for transport-properties of inhomogeneous porous rock. *Journal of Geophysical Research, Solid Earth* 97: 17409–17424.

Berryman JG (1993) Effective-stress rules for pore-fluid transport in rocks containing 2 minerals. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* 30: 1165–1168.

Berryman JG (1995) Mixture theories for rock properties. In: *Rock Physics & Phase Relations*, pp. 205–228. Washington, DC: American Geophysical Union.

Berryman JG (1999) Origin of gassmann's equations. *Geophysics* 64: 1627–1629.

Berryman JG (2005) Bounds and self-consistent estimates for elastic constants of random polycrystals with hexagonal, trigonal, and tetragonal symmetries. *Journal of the Mechanics and Physics of Solids* 53: 2141–2173.

Berryman JG (2012) Evaluating bounds and estimators for constants of random polycrystals composed of orthotropic elastic materials. *International Journal of Engineering Science* 58: 11–20.

Berryman JG and Berge PA (1996) Critique of two explicit schemes for estimating elastic properties of multiphase composites. *Mechanics of Materials* 22: 149–164.

Berryman JG, Pride SR, and Wang HF (2002) A differential scheme for elastic properties of rocks with dry or saturated cracks. *Geophysical Journal International* 151: 597–611.

Bett KE, Rowlinson JS, and Saville G (2003) *Thermodynamics for Chemical Engineers*. Cambridge: MIT Press.

Bianco E, Kaplan S, and Schmitt DR (2010) Seismic rock physics of steam injection in bituminous-oil reservoirs – Chapter 6. In: Batzle ML, Chopra S, Lines LR, and Schmitt DR (eds.) *Heavy Oils: Reservoir Characterization and Production Monitoring*, pp. 105–110. Tulsa: Society of Exploration Geophysicists.

Biot MA (1956a) Theory of propagation of elastic waves in a fluid-saturated porous solid.1. Low-frequency range. *Journal of the Acoustical Society of America* 28: 168–178.

Biot MA (1956b) Theory of propagation of elastic waves in a fluid-saturated porous solid.2. Higher frequency range. *Journal of the Acoustical Society of America* 28: 179–191.

Biot MA and Willis D (1957) The elastic coefficients of the theory of consolidation. *Journal of Applied Mechanics-Transactions of the ASME* 24: 594–601.

Birch F (1947) Finite elastic strain of cubic crystals. *Physical Review* 71: 809–824.

Birch F (1960) The velocity of compressional waves in rocks to 10-kilobars.1. *Journal of Geophysical Research* 65: 1083–1102.

Birch F (1961) The velocity of compressional waves in rocks to 10 kilobars, part 2. *Journal of Geophysical Research* 66: 2199–2224.

Blake OO, Faulkner DR, and Rietbrock A (2013) The effect of varying damage history in crystalline rocks on the *P*- and *S*-wave velocity under hydrostatic confining pressure. *Pure and Applied Geophysics* 170: 493–505.

Bonner BP (1974) Shear wave birefringence in dilating granite. *Geophysical Research Letters* 1: 217–220.

Bourbié T, Coussy O, and Zinszner B (1987) *Acoustics of Porous Media*. Houston, TX: Gulf.

Bouzidi YS and Schmitt DR (2009) Measurement of the speed and attenuation of the biot slow wave using a large ultrasonic transmitter. *Journal of Geophysical Research* 114: B08201.

Bower AF (2010) *Applied Mechanics of Solids*. Boca Raton: Taylor and Francis.

Brace WF (1965) Some new measurements of linear compressibility of rocks. *Journal of Geophysical Research* 70: 391–398.

Brace WF and Martin RJ (1968) A test of law of effective stress for crystalline rocks of low porosity. *International Journal of Rock Mechanics and Mining Sciences* 5: 415–426.

Brown JM, Abramson EH, and Angel RJ (2006) Triclinic elastic constants for low albite. *Physics and Chemistry of Minerals* 33: 256–265.

Budiansky B and O'Connell RJ (1976) Elastic-moduli of a cracked solid. *International Journal of Solids and Structures* 12: 81–97.

Bunge H (1974) The effective elastic constants of textured polycrystals in second order approximation. *Kristall und Technik* 9: 413–424.

Calderon E, Gauthier M, Decremps F, Hamel G, Syfosse G, and Polian A (2007) Complete determination of the elastic moduli of alpha-quartz under hydrostatic pressure up to 1 GPa: An ultrasonic study. *Journal of Physics: Condensed Matter* 19: 436228.

Carcione JM (2007) *Wave Fields in Real Media: Wave Propagation in Anisotropic, Anelastic, Porous and Electromagnetic Media*. Amsterdam: Elsevier Science.

Carcione JM, Picotti S, Gei D, and Rossi G (2006) Physics and seismic modeling for monitoring CO_2 storage. *Pure and Applied Geophysics* 163: 175–207.

Carcione JM and Tinivella U (2001) The seismic response to overpressure: A modelling study based on laboratory, well and seismic data. *Geophysical Prospecting* 49: 523–539.

Carey WM and Evans RB (2011) *A Review of the Sonic Properties of Bubbly Liquids*. New York: Springer.

Carvalho FCS and Labuz JF (1996) Experiments on effective elastic modulus of two-dimensional solids with cracks and holes. *International Journal of Solids and Structures* 33: 4119–4130.

Castagna JP, Batzle ML, and Eastwood RL (1985) Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. *Geophysics* 50: 571–581.

Castagna JP, Batzle ML, and Kan TK (1993) Rock physics – The link between rock properties and AVO response. In: Castagna JP and Backus M (eds.) *Offset-Dependent Reflectivity – Theory and Practice of AVO Analysis*, pp. 135–171. Tulsa: Society of Exploration Geophysicists.

Castellan GW (1971) *Physical Chemistry*, 2nd edn. New York: Addison-Wesley.

Castier M (2011) Thermodynamic speed of sound in multiphase systems. *Fluid Phase Equilibria* 306: 204–211.

Chen C-T and Millero F (1976) Reevaluation of Wilson's sound-speed measurements for pure water. *The Journal of the Acoustical Society of America* 60: 1270–1273.

Cheng C and Toksöz M (1979) Inversion of seismic velocities for the pore aspect ratio spectrum of a rock. *Journal of Geophysical Research, Solid Earth* 84: 7533–7543.

Cheyne SA, Stebbings CT, and Roy RA (1995) Phase-velocity measurements in bubbly liquids using a fiber optic laser interferometer. *Journal of the Acoustical Society of America* 97: 1621–1624.

Chitale DV and Sigal R (2000) NMR characterization of the water adsorbed by montmorillonite: Impact on the analysis of porosity logs: Spe 62531. In: *SPE/AAPG Western Regional Meeting*, Long Beach: Society of Petroleum Engineers.

Cholach PY, Molyneux JB, and Schmitt DR (2005) Flin flon belt seismic anisotropy: Elastic symmetry, heterogeneity, and shear-wave splitting. *Canadian Journal of Earth Sciences* 42: 533–554.

Cholach PY and Schmitt DR (2006) Intrinsic elasticity of a textured transversely isotropic muscovite aggregate: Comparisons to the seismic anisotropy of schists and shales. *Journal of Geophysical Research, Solid Earth* 30: 111–116.

Christensen NI (1982) Seismic velocities. In: Carmichael RS (ed.) *Handbook of Physical Properties of Rocks*, vol. II, pp. 1–228. Boca Raton: CRC Press.

Christensen NI (1989) Seismic velocities. In: Carmichael R (ed.) *Practical Handbook of Physical Properties of Rocks and Minerals*, pp. 429–546. Boca Raton: CRC.

Christensen NI (2004) Serpentinites, peridotites, and seismology. *International Geology Review* 46: 795–816.

Christensen NI and Mooney WD (1995) Seismic velocity structure and composition of the continental-crust – A global view. *Journal of Geophysical Research, Solid Earth* 100: 9761–9788.

Christensen NI and Stanley D (2003) Seismic velocities and densities of rocks. In: Lee W, Kanamori H, Jennings P, and Kissinger C (eds.) *International Handbook of Earthquake and Engineering Seismology*, pp. 1587–1594. New York: Academic Press.

Christensen NI and Wang HF (1985) The influence of pore pressure and confining pressure on dynamic elastic properties of berea sandstone. *Geophysics* 50: 207–213.

Ciz R and Shapiro SA (2009) Stress-dependent anisotropy in transversely isotropic rocks: Comparison between theory and laboratory experiment on shale. *Geophysics* 74: D7–D12.

Clark VA (1992) The effect of oil under in situ conditions on the seismic properties of rocks. *Geophysics* 57: 894–901.

Commander KW and Prosperetti A (1989) Linear pressure waves in bubbly liquids – Comparison between theory and experiments. *Journal of the Acoustical Society of America* 85: 732–746.

Condon E (1933) Note on the velocity of sound. *American Journal of Physics* 1: 18.

Coppens A (1981) Simple equations for the speed of sound in neptunian waters. *The Journal of the Acoustical Society of America* 69: 862–863.

Crampin S (1994) The fracture criticality of crustal rocks. *Geophysical Journal International* 118: 428–438.

Crampin S and Peacock S (2008) A review of the current understanding of seismic shear-wave splitting in the earth's crust and common fallacies in interpretation. *Wave Motion* 45: 675–722.

Crosson R and Lin J (1971) Voigt and reuss prediction of anisotropic elasticity of dunite. *Journal of Geophysical Research* 76: 570–578.

Cundall PA and Strack ODL (1979) Discrete numerical-model for granular assemblies. *Geotechnique* 29: 47–65.

Daridon JL, Lagourette B, Xans P, and Montel F (1998) Petroleum characterization from ultrasonic measurement. *Journal of Petroleum Science and Engineering* 19: 281–293.

David EC and Zimmerman RW (2011) Elastic moduli of solids containing spheroidal pores. *International Journal of Engineering Science* 49: 544–560.

David EC and Zimmerman RW (2012) Pore structure model for elastic wave velocities in fluid-saturated sandstones. *Journal of Geophysical Research, Solid Earth* 117: B07210.

Del Grosso VA (1974) New equation for speed of sound in natural-waters (with comparisons to other equations). *Journal of the Acoustical Society of America* 56: 1084–1091.

Del Grosso VA and Mader CW (1972) Speed of sound in pure water. *Journal of the Acoustical Society of America* 52: 1442–1446.

Diamond LW and Akinfiev NN (2003) Solubility of CO_2 in water from −1.5 to 100 degrees c and from 0.1 to 100 MPa: Evaluation of literature data and thermodynamic modelling. *Fluid Phase Equilibria* 208: 265–290.

Digby PJ (1981) The effective elastic-moduli of porous granular rocks. *Journal of Applied Mechanics-Transactions of the ASME* 48: 803–808.

Dillen MWP, Cruts HMA, Groenenboom J, Fokkema JT, and Duijndam AJW (1999) Ultrasonic velocity and shear-wave splitting behavior of a colton sandstone under a changing triaxial stress. *Geophysics* 64: 1603–1607.

Domenico S (1977) Elastic properties of unconsolidated porous sand reservoirs. *Geophysics* 42: 1339–1368.

Duan Z, Hu J, Li D, and Mao S (2008) Densities of the CO_2–H_2O and CO_2–H_2O–NaCl systems up to 647 K and 100 MPa. *Energy & Fuels* 22: 1666–1674.

Duan ZH and Sun R (2003) An improved model calculating CO_2 solubility in pure water and aqueous NaCl solutions from 273 to 533 k and from 0 to 2000 bar. *Chemical Geology* 193: 257–271.

Dvorkin JP (2008) Yet another Vs equation. *Geophysics* 73: E35–E39.

Dvorkin J, Mavko G, and Nur A (1995) Squirt flow in fully saturated rocks. *Geophysics* 60: 97–107.

Eberhart-Phillips D, Han DH, and Zoback MD (1989) Empirical relationships among seismic velocity, effective pressure, porosity, and clay content in sandstone. *Geophysics* 54: 82–89.

Ebrahimi D, Pellenq RJM, and Whittle AJ (2012) Nanoscale elastic properties of montmorillonite upon water adsorption. *Langmuir* 28: 16855–16863.

Elbra T, Karlqvist R, Lassila I, Haeggstrom E, and Pesonen LJ (2011) Laboratory measurements of the seismic velocities and other petrophysical properties of the Outokumpu deep drill core samples, eastern Finland. *Geophysical Journal International* 184: 405–415.

Evelein K, Moore R, and Heidemann R (1976) Correlation of phase behavior in systems hydrogen-sulfide – Water and carbon-dioxide–water. *Industrial & Engineering Chemistry Process Design and Development* 15: 423–428.

Ezekwe N (2010) *Petroleum Reservoir Engineering Practice*. New York: Prentice Hall.

Fabricius IL, Gommesen L, Krogsboll A, and Olsen D (2008) Chalk porosity and sonic velocity versus burial depth: Influence of fluid pressure, hydrocarbons, and mineralogy. *AAPG Bulletin* 92: 201–223.

Fam M and Santamarina JC (1997) A study of consolidation using mechanical and electromagnetic waves. *Geotechnique* 47: 203–219.

Fereidoon V, Embleton J, and Lines LR (2010) The effects of cold production on seismic response. In: Chopra S, Lines LR, Schmitt DR, and Batzle M (eds.) *Heavy Oils: Reservoir Characterization and Production Monitoring*, pp. 237–241. Tulsa: Society of Exploration Geophysicists, ch. 18.

Fletcher NH (1974) Adiabatic assumption for wave-propagation. *American Journal of Physics* 42: 487–489.

Fletcher N (1976) Adiabatic assumption for wave propagation. *American Journal of Physics* 44: 486–487.

Freund D (1992) Ultrasonic compressional and shear velocities in dry clastic rocks as a function of porosity, clay content, and confining pressure. *Geophysical Journal International* 108: 125–135.

Fridleifsson GO and Elders WA (2005) The iceland deep drilling project: A search for deep unconventional geothermal resources. *Geothermics* 34: 269–285.

Fuck RF and Tsvankin I (2009) Analysis of the symmetry of a stressed medium using nonlinear elasticity. *Geophysics* 74: WB79–WB87.

Fung YC (1965) *Foundations of Solid Mechanics*. New York: Prentice Hall.

Gallagher JS, Crovetto R, and Sengers J (1993) The thermodynamic behavior of the CO_2–H_2O system from 400-K to 1000-K, up to 100 MPa and 30-percent mole fraction of CO_2. *Journal of Physical and Chemical Reference Data* 22: 431–513.

Gallop J (2013) Extended walton third-order elastic coefficients modified by an anisotropic and stress-dependent coordination number. *Geophysics* 78: D545–D556.

Gao K and Gibson RL Jr. (2012) Pressure-dependent seismic velocities based on effective compliance theory and an asperity deformation model. *Geophysics* 77: D229–D243.

Gardner GHF, Gardner LW, and Gregory AR (1974) Formation velocity and density – Diagnostic basics for stratigraphic traps. *Geophysics* 39: 770–780.

Gassmann F (1951) Über die elastizität poröser medien. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zurich* 96: 1–23.

Gebrande H (1982) Elastic wave velocities and constants of elasticity at normal conditions. In: *The Landolt-Bornstein Database*, pp. 1–99. Berlin: Springer-Verlag.

Gercek H (2007) Poisson's ratio values for rocks. *International Journal of Rock Mechanics and Mining Sciences* 44: 1–13.

Goddard JD (1990) Nonlinear elasticity and pressure-dependent wave speeds in granular media. *Proceedings of The Royal Society A Mathematical Physical and Engineering Sciences* 430: 105–131.

Gomez CT, Dvorkin J, and Vanorio T (2010) Laboratory measurements of porosity, permeability, resistivity, and velocity on fontainebleau sandstones. *Geophysics* 75: E191–E204.

Gray C, Gubbins K, and Joslin C (2011) Thermodynamic properties of mixtures. In: Gray C, Gubbins K, and Joslin C (eds.) *Theory of Molecular Fluids. Applications*, vol. 2, pp. 761–886. New York: Oxford University Press.

Greenberg ML and Castagna JP (1992) Shear-wave velocity estimation in porous rocks – Theoretical formulation, preliminary verification and applications. *Geophysical Prospecting* 40: 195–209.

Greenberg ML and Castagna JP (1993) Shear-wave velocity estimation in porous rocks – Theoretical formulation, preliminary verification and applications (vol 40, pg 195, 1992). *Geophysical Prospecting* 41: 515.

Gucker FT, Stubley D, and Hill DJ (1975) Isentropic compressibilities of aqueous-solutions of some alkali-halides at 298.15 K. *Journal of Chemical Thermodynamics* 7: 865–873.

Gueguen Y and Bouteca M (2004) *Mechanics of Fluid-Saturated Rocks*. Amsterdam/Boston: Elsevier/Academic Press.

Gurevich B, Pervukhina M, and Makarynska D (2011) An analytic model for the stress-induced anisotropy of dry rocks. *Geophysics* 76: WA125–WA133.

Hamilton EL (1979) $V_p - V_s$ and Poisson ratios in marine-sediments and rocks. *Journal of the Acoustical Society of America* 66: 1093–1101.

Han DH (1986) *Effects of Porosity and Clay Content on Acoustic Properties of Sandstones and Unconsolidated Sediments*. PhD, Stanford University.

Han D-H and Batzle M (2000) Velocity, density and modulus of hydrocarbon fluids – Data measurement. In: *Society of Exploration Geophysicists 70th Annual Meeting and Exposition, Calgary: Society of Exploration Geophysicist*.

Han DH and Batzle ML (2004) Gassmann's equation and fluid-saturation effects on seismic velocities. *Geophysics* 69: 398–405.

Han D-H, Sun M, and Batzle M (2010) CO_2 velocity measurement and models for temperatures up to 200 degrees C and pressures up to 100 MPa. *Geophysics* 75: E123–E129.

Han D-H, Sun M, and Batzle M (2011) CO_2 velocity measurement and models for temperatures up to 200 degrees C and pressures up to 100 MPa. *Geophysics* 75: e123–e129.

Hardin BO and Blandford GE (1989) Elasticity of particulate materials. *Journal of Geotechnical Engineering, ASCE* 115: 788–805.

Hashin Z and Shtrikman S (1962) On some variational principles in anisotropic and nonhomogeneous elasticity. *Journal of the Mechanics and Physics of Solids* 10: 335–342.

Hashin Z and Shtrikman S (1963) A variational approach to the theory of the elastic behaviour of multiphase materials. *Journal of the Mechanics and Physics of Solids* 11: 127–140.

He T (2006) *P and S Wave Velocity Measurement and Pressure Sensitivity Analysis of AVA Response*. MSc, University of Alberta.

Hearmon RFS (1946) The elastic constants of anisotropic materials. *Reviews of Modern Physics* 18: 409–440.

Hearmon RFS (1956) The elastic constants of anisotropic materials – II. *Advances in Physics* 5: 323–382.

Helgeson HC and Kirkham DH (1974) Theoretical prediction of thermodynamic behavior of aqueous electrolytes at high pressures and temperatures. 1. Summary of thermodynamic-electrostatic properties of solvent. *American Journal of Science* 274: 1089–1198.

Hemley RJ, Mao HK, and Struzhkin VV (2005) Synchrotron radiation and high pressure: New light on materials under extreme conditions. *Journal of Synchrotron Radiation* 12: 135–154.

Herwanger J and Horne S (2005) Predicting time-lapse stress effects in seismic data. *The Leading Edge* 24: 1234–1242.

Hill R (1952) The elastic behaviour of a crystalline aggregate. *Proceedings of the Physical Society. Section A* 65: 349–355.

Hill R (1965) A self-consistent mechanics of composite materials. *Journal of the Mechanics and Physics of Solids* 13: 213–222.

Holland TJB and Powell R (1998) An internally consistent thermodynamic data set for phases of petrological interest. *Journal of Metamorphic Geology* 16: 309–343.

Holmes GM, Crampin S, and Young RP (2000a) Controlled sources for shear-wave surveys in mines. *Geophysical Prospecting* 48: 399–414.

Holmes GM, Crampin S, and Young RP (2000b) Seismic anisotropy in granite at the underground research laboratory, Manitoba. *Geophysical Prospecting* 48: 415–435.

Holzapfel GA (2000) *Nonlinear Solid Mechanics: A Continuum Approach for Engineering*. Chichester, NY: Wiley.

Horii H and Nemat-Nasser S (1983) Overall moduli of solids with microcracks: Load-induced anisotropy. *Journal of the Mechanics and Physics of Solids* 31: 155–171.

Hou SX, Maitland GC, and Trusler JPM (2013) Measurement and modeling of the phase behavior of the (carbon dioxide plus water) mixture at temperatures from 298.15 K to 448.15 K. *Journal of Supercritical Fluids* 73: 87–96.

Hu J, Duan Z, Zhu C, and Chou IM (2007) Pvtx properties of the CO_2–H_2O and CO_2–H_2O–NaCl systems below 647 K: Assessment of experimental data and thermodynamic models. *Chemical Geology* 238: 249–267.

Hu M, O'sullivan C, Jardine R, and Jiang M (2010) Stress-induced anisotropy in sand under cyclic loading. *Granular Matter* 12: 469–476.

Hudson JA (1981) Wave speeds and attenuation of elastic-waves in material containing cracks. *Geophysical Journal of the Royal Astronomical Society* 64: 133–150.

Huntington HB (1958) The elastic constants of crystals. *Solid State Physics: Advances in Research and Applications* 7: 213–351.

IAPWS (2008) *Release on the IAPWS Formulation 2008 for the Thermodynamic Properties of Seawater*. Berlin: International Association for the Properties of Water and Steam.

IAPWS (2011) *Revised Release on the Pressure along the Melting and Sublimation Curves of Ordinary Water Substance*, Available: http://www.iapws.org/relguide/MeltSub2011.pdf (accessed 25 June 2013).

Ji SC, Wang Q, and Xia B (2002) *Handbook of Seismic Properties of Minerals, Rocks and Ores*. Montreal: Polytechnic International Press.

Johnson DL (1984) Recent developments in the acoustic properties of porous media. In: Sette D (ed.) *Frontiers in Physical Acoustics*, *Proceedings of the International School of Physics "Enrico Fermi" Course XCIII*, pp. 255–290.

Johnson K (1987) *Contact Mechanics*. New York: Cambridge University Press.

Johnson PA and Rasolofosaon PNJ (1996) Nonlinear elasticity and stress-induced anisotropy in rock. *Journal of Geophysical Research, Solid Earth* 101: 3113–3124.

Jones T and Nur A (1983) Velocity and attenuation in sandstone at elevated-temperatures and pressures. *Geophysical Research Letters* 10: 140–143.

Jorstad A, Mukerji T, and Mavko G (1999) Model-based shear-wave velocity estimation versus empirical regressions. *Geophysical Prospecting* 47: 785–797.

Kachanov M (1980) Continuum model of medium with cracks. *Journal of the Engineering Mechanics Division, ASCE* 106: 1039–1051.

Kachanov M, Tsukrov I, and Shafiro B (1994) Effective moduli of solids with cavities of various shapes. *Applied Mechanics Reviews* 1: S151–S174.

Kanitpanyacharoen W, Wenk HR, Kets F, Lehr C, and Wirth R (2011) Texture and anisotropy analysis of qusaiba shales. *Geophysical Prospecting* 59: 536–556.

Karmous MS (2011) Theoretical study of kaolinite structure; energy minimization and crystal properties. *The World Journal of Nano Science and Engineering* 1: 62–66.

Kay WB (1936) Density of hydrocarbon gases and vapors at high temperature and pressure. *Industrial and Engineering Chemistry* 28: 1014–1019.

Kell GS (1975) Density, thermal expansivity, and compressibility of liquid water from 0 degrees to 150 degrees. Correlations and tables for atmospheric-pressure and saturation reviewed and expressed on 1968 temperature scale. *Journal of Chemical and Engineering Data* 20: 97–105.

Kern H (1982) Elastic wave velocities and constants of elasticity of rocks at elevated pressures and temperatures. In: *The Landolt-Bornstein Database*, pp. 99–140. Berlin: Springer-Verlag.

Kern H, Mengel K, Strauss KW, Ivankina TI, Nikitin AN, and Kukkonen IT (2009) Elastic wave velocities, chemistry and modal mineralogy of crustal rocks sampled by the outokumpu scientific drill hole: Evidence from lab measurements and modeling. *Physics of the Earth and Planetary Interiors* 175: 151–166.

Khaksar A, Griffiths CM, and Mccann C (1999) Compressional- and shear-wave velocities as a function of confining stress in dry sandstones. *Geophysical Prospecting* 47: 487–508.

Khazanehdari J and Mccann C (2005) Acoustic and petrophysical relationships in low-shale sandstone reservoir rocks. *Geophysical Prospecting* 53: 447–461.

Khazanehdari J and Sothcott J (2003) Variation in dynamic elastic shear modulus of sandstone upon fluid saturation and substitution. *Geophysics* 68: 472–481.

Khidas Y and Jia XP (2010) Anisotropic nonlinear elasticity in a spherical-bead pack: Influence of the fabric anisotropy. *Physical Review E* 81: 021303.

Kieffer SW (1977) Sound speed in liquid-gas mixtures – Water-air and water-steam. *Journal of Geophysical Research* 82: 2895–2904.

Kirstetter O, Corbett P, Somerville J, and Macbeth C (2006) Elasticity/saturation relationships using flow simulation from an outcrop analogue for 4D seismic modelling. *Petroleum Geoscience* 12: 205–219.

Kumar A (2003) Speed of sound in concentrated aqueous KCl solutions from 278.15 to 338.15 K. *Journal of Chemical and Engineering Data* 48: 388–391.

Kumazawa M (1969) The elastic constant of polycrystalline rocks and nonelastic behavior inherent to them. *Journal of Geophysical Research* 74: 5311–5320.

Kunz O and Wagner W (2012) The gerg-2008 wide-range equation of state for natural gases and other mixtures: An expansion of gerg-2004. *Journal of Chemical and Engineering Data* 57: 3032–3091.

Kuster GT and Toksöz MN (1974a) Velocity and attenuation of seismic-waves in 2-phase media. 1. Theoretical formulations. *Geophysics* 39: 587–606.
Kuster GT and Toksöz MN (1974b) Velocity and attenuation of seismic-waves in 2-phase media. 2. Experimental results. *Geophysics* 39: 607–618.
Kwon O, Kronenberg AK, Gangi AF, and Johnson B (2001) Permeability of wilcox shale and its effective pressure law. *Journal of Geophysical Research, Solid Earth* 106: 19339–19353.
Lakes RS (2009) *Viscoelastic Materials*. New York: Cambridge University Press.
Laliberte M and Cooper WE (2004) Model for calculating the density of aqueous electrolyte solutions. *Journal of Chemical and Engineering Data* 49: 1141–1151.
Landau LD and Lifshitz EM (1959) *Fluid Mechanics*. New York: Pergamon Press.
Landau LD and Lifshitz EM (1970) *Theory of Elasticity*. Oxford: Pergamon Press.
Lassila I, Karlqvist R, Elbra T, Gates FK, Pesonen LJ, and Haeggstrom E (2010) Ultrasonic velocity of the upper gneiss series rocks from the outokumpu deep drill hole, fennoscandian shield – Comparing uniaxial to triaxial loading. *Journal of Applied Geophysics* 72: 178–183.
Lemmon EW, Jacobsen RT, Penoncello SG, and Friend DG (2000) Thermodynamic properties of air and mixtures of nitrogen, argon, and oxygen from 60 to 2000 k at pressures to 2000 MPa. *Journal of Physical and Chemical Reference Data* 29: 331–385.
Lemmon EW, Mclinden MO, and Friend DG (2012a) Thermophysical properties of fluid systems. In: Lindstrom PJ and Mallard WG (eds.) *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. Gaithersburg, MD: National Institute of Standards and Technology.
Lemmon EW, Mclinden MO, and Friend DG (eds.) (2012b) *Thermophysical Properties of Fluid Systems*. Gaithersburg, MD: National Institute of Standards and Technology.
Leravalec M and Gueguen Y (1996) High- and low-frequency elastic moduli for a saturated porous/cracked rock – Differential self-consistent and poroelastic theories. *Geophysics* 61: 1080–1094.
Li HB and Zhang JJ (2011) Elastic moduli of dry rocks containing spheroidal pores based on differential effective medium theory. *Journal of Applied Geophysics* 75: 671–678.
Lin C-W and Trusler JPM (2012) The speed of sound and derived thermodynamic properties of pure water at temperatures between (253 and 473) k and at pressures up to 400 MPa. *Journal of Chemical Physics* 136: 094511.
Lo TW, Coyner KB, and Toksöz MN (1986) Experimental-determination of elastic-anisotropy of berea sandstone, chicopee shale, and chelmsford granite. *Geophysics* 51: 164–171.
Lowell S, Shields JE, Thomas MA, and Thommes M (2004) *Characterization of Porous Solids and Powders: Surface Area, Pore Size and Density*. Amsterdam: Springer.
Lu C and Jackson I (2006) Low-frequency seismic properties of thermally cracked and argon-saturated granite. *Geophysics* 71: F147–F159.
Lumley DE (2001) Time-lapse seismic reservoir monitoring. *Geophysics* 66: 50–53.
Lyakhovsky V, Reches Z, Weinberger R, and Scott TE (1997) Non-linear elastic behaviour of damaged rocks. *Geophysical Journal International* 130: 157–166.
Mackenzie JK (1950) The elastic constants of a solid containing spherical holes. *Proceedings of the Physical Society. Section B* 63: 2–11.
Mackenzie KV (1981) 9-term equation for sound speed in the oceans. *Journal of the Acoustical Society of America* 70: 807–812.
Mainprice D (1990) A fortran program to calculate seismic anisotropy from the lattice preferred orientation of minerals. *Computers & Geosciences* 16: 385–393.
Mainprice D and Humbert M (1994) Methods of calculating petrophysical properties from lattice preferred orientation data. *Surveys in Geophysics* 15: 575–592.
Makse HA, Gland N, Johnson DL, and Schwartz LM (1999) Why effective medium theory fails in granular materials. *Physical Review Letters* 83: 5070–5073.
Man CS and Huang MJ (2011) A simple explicit formula for the Voigt–Reuss–Hill average of elastic polycrystals with arbitrary crystal and texture symmetries. *Journal of Elasticity* 105: 29–48.
Mao S, Duan Z, and Hu W (2009) A vapor-liquid phase equilibrium model for binary CO_2–H_2O and CH_4–H_2O systems above 523 K for application to fluid inclusions. *The Journal of Supercritical Fluids* 50: 13–21.
Matthies S and Humbert M (1993) The realization of the concept of a geometric mean for calculating physical constants of polycrystalline materials. *Physica Status Solidi B: Basic Research* 177: K47–K50.
Mavko G and Jizba D (1991) Estimating grain-scale fluid effects on velocity dispersion in rocks. *Geophysics* 56: 1940–1949.
Mavko G, Mukerji T, and Dvorkin J (2003) *The Rock Physics Handbook: Tools for Seismic Analysis in Porous Media*. New York: Cambridge University Press.
Mavko G and Nur A (1975) Melt squirt in asthenosphere. *Journal of Geophysical Research* 80: 1444–1448.
Mavko GM and Nur A (1978) The effect of nonelliptical cracks on the compressibility of rocks. *Journal of Geophysical Research, Solid Earth* 83: 4459–4468.
Mavko G and Vanorio T (2010) The influence of pore fluids and frequency on apparent effective stress behavior of seismic velocities. *Geophysics* 75: N1–N7.
Meister R and Peselnick L (1966) Variational method of determining effective moduli of polycrystals with tetragonal symmetry. *Journal of Applied Physics* 37: 4121–4125.
Melendez Martinez J and Schmitt DR (2013) Anisotropic elastic moduli of carbonates and evaporites from the Weyburn-Midale reservoir and seal rocks. *Geophysical Prospecting* 61: 363–379.
Militzer B, Wenk H-R, Stackhouse S, and Stixrude L (2011) First-principles calculation of the elastic moduli of sheet silicates and their application to shale anisotropy. *American Mineralogist* 96: 125–137.
Mookherjee M and Stixrude L (2009) Structure and elasticity of serpentine at high-pressure. *Earth and Planetary Science Letters* 279: 11–19.
Mori T and Tanaka K (1973) Average stress in matrix and average elastic energy of materials with misfitting inclusions. *Acta Metallurgica* 21: 571–574.
Morris PR (2006) Polycrystal elastic constants for triclinic crystal and physical symmetry. *Journal of Applied Crystallography* 39: 502–508.
Nichita DV, Khalid P, and Broseta D (2010) Calculation of isentropic compressibility and sound velocity in two-phase fluids. *Fluid Phase Equilibria* 291: 95–102.
Nist (2010a) *NIST Reference Fluid Thermodynamic and Transport Properties Database (RefProp): Version 9.1 [Online]*. Gaithersburg, MD: National Institute of Standards and Technology, Available: http://www.nist.gov/srd/nist23.cfm.
Nist (2010b) *NIST/ASME Steam Properties Database: Version 3.0, the Latest IAPWS Standard Formulation [Online]*. Gaithersburg, MD: National Institute of Standards and Technology, Available: http://www.nist.gov/srd/nist10.cfm.
Njiekak G, Schmitt DR, Yam H, and Kofman RS (2013) CO_2 rock physics as part of the Weyburn-Midale geological storage project. *International Journal of Greenhouse Gas Control* 16S: S118–S133.
Norris AN (1985) A differential scheme for the effective moduli of composites. *Mechanics of Materials* 4: 1–16.
Nur A (1971) Effects of stress on velocity anisotropy in rocks with cracks. *Journal of Geophysical Research* 76: 2022–2034.
Nur A and Byerlee JD (1971) An exact effective stress law for elastic deformation of rock with fluids. *Journal of Geophysical Research* 76: 6414–6419.
Nur A and Simmons G (1969) Stress-induced velocity anisotropy in rock – An experimental study. *Journal of Geophysical Research* 74: 6667–6674.
O'Connell RJ and Budiansky B (1977) Viscoelastic properties of fluid-saturated cracked solids. *Journal of Geophysical Research* 82: 5719–5735.
Oakley BA, Barber G, Worden T, and Hanna D (2003a) Ultrasonic parameters as a function of absolute hydrostatic pressure. I. A review of the data for organic liquids. *Journal of Physical and Chemical Reference Data* 32: 1501–1533.
Oakley BA, Hanna D, Shillor M, and Barber G (2003b) Ultrasonic parameters as a function of absolute hydrostatic pressure. II. Mathematical models of the speed of sound in organic liquids. *Journal of Physical and Chemical Reference Data* 32: 1535–1544.
O'Connell RJ and Budiansky B (1974) Seismic velocities in dry and saturated cracked solids. *Journal of Geophysical Research* 79: 5412–5426.
Ohno I, Abe M, Kimura M, Hanayama Y, Oda H, and Suzuki I (2000) Elasticity measurement of silica glass under gas pressure. *American Mineralogist* 85: 288–291.
Ohno I, Harada K, and Yoshitomi C (2006) Temperature variation of elastic constants of quartz across the alpha-beta transition. *Physics and Chemistry of Minerals* 33: 1–9.
Payan C, Garnier V, Moysan J, and Johnson PA (2009) Determination of third order elastic constants in a complex solid applying coda wave interferometry. *Applied Physics Letters* 94: 011904.
Peng DY and Robinson DB (1976) A new two-constant equation of state. *Industrial and Engineering Chemistry Fundamentals* 15: 59–64.
Peselnick L and Meister R (1965) Variational method of determining effective moduli of polycrystals – (a) Hexagonal symmetry (b) trigonal symmetry. *Journal of Applied Physics* 36: 2879–2884.
Pham DC (2011) On the scatter ranges for the elastic moduli of random aggregates of general anisotropic crystals. *Philosophical Magazine* 91: 609–627.
Picard DJ and Bishnoi PR (1987) Calculation of the thermodynamic sound-velocity in 2-phase multicomponent fluids. *International Journal of Multiphase Flow* 13: 295–308.
Pickett GR (1963) Acoustic character logs and their applications in formation evaluation. *Transactions of the Society of Petroleum Engineers of AIME* 228: 659–667.
Plantier F, Bessieres D, Daridon JL, and Montel F (2008) Structure and thermodynamic consistency of heavy oils: A study on the basis of acoustic measurements. *Fuel* 87: 196–201.

Plona TJ (1980) Observation of a 2nd bulk compressional wave in a porous-medium at ultrasonic frequencies. *Applied Physics Letters* 36: 259–261.

Prasad M and Manghnani MH (1997) Effects of pore and differential pressure on compressional wave velocity and quality factor in berea and michigan sandstones. *Geophysics* 62: 1163–1176.

Prikryl R, Klima K, Lokajicek T, and Pros Z (2005) Non-linearity in multidirectional *P*-wave velocity; confining pressure behaviour based on real 3D laboratory measurements, and its mathematical approximation. In: Brewer TS, Pezard PA, Petrov VA, and Harvey PK (eds.) *Petrophysical Properties of Crystalline Rocks*, pp. 323–334. London: Geological Society of London.

Prioul R, Bakulin A, and Bakulin V (2004) Nonlinear rock physics model for estimation of 3d subsurface stress in anisotropic formations: Theory and laboratory verification. *Geophysics* 69: 415–425.

Rai CS and Hanson KE (1988) Shear-wave velocity anisotropy in sedimentary-rocks – A laboratory study. *Geophysics* 53: 800–806.

Raymer L, Hunt ER, and Gardner J (1980) An improved sonic transit time-to-porosity transform. In: Paper P (ed.) *21st Annual Logging Symposium*, Layfayette: Soc. of Well Log Analysts.

Reuss A (1929a) Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle. *Journal of Applied Mathematics and Mechanics (Z. Angew. Math. Mech)* 9: 49–58.

Reuss A (1929b) Berechnungd er fliessgrenzey on mischkristallen auf grund der plastiziffitsbedingunfagr einkristall. *Zeitschrift für Angewandte Mathematik und Mechanik* 9: 49–58.

Rice JR and Cleary MP (1976) Some basic stress diffusion solutions for fluid-saturated elastic porous-media with compressible constituents. *Reviews of Geophysics* 14: 227–241.

Rivers ML and Carmichael ISE (1987) Ultrasonic studies of silicate melts. *Journal of Geophysical Research-Solid Earth and Planets* 92: 9247–9270.

Roesler SK (1979) Anisotropic shear modulus due to stress anisotropy. *Journal of the Geotechnical Engineering Division, ASCE* 105: 871–880.

Rogers PSZ and Pitzer KS (1982) Volumetric properties of aqueous sodium-chloride solutions. *Journal of Physical and Chemical Reference Data* 11: 15–81.

Rowe AM and Chou JCS (1970) Pressure–volume–temperature–concentration relation of aqueous NaCl solutions. *Journal of Chemical and Engineering Data* 15: 61–66.

Ryzhova TV, Alexandrov KS, and Korobkova VM (1966) The elastic properties of rock forming minerals v. Additional data on silicates. *Izvestiya, Earth Physics* 2: 63–65.

Safarov J, Berndt S, Millero F, Feistel R, Heintz A, and Hassel E (2012) (p, ρ, t) properties of seawater: Extensions to high salinities. *Deep Sea Research Part I: Oceanographic Research Papers* 65: 146–156.

Safarov J, Berndt S, Millero FJ, Feistel R, Heintz A, and Hassel E (2013) (p, r, t) properties of seawater at brackish salinities: Extension to high temperatures and pressures. *Deep-Sea Research Part 1* 78: 95–101.

Safarov J, Millero F, Feistel R, Heintz A, and Hassel E (2009) Thermodynamic properties of standard seawater: Extensions to high temperatures and pressures. *Ocean Science* 5: 235–246.

Sato H, Ono K, Johnston CT, and Yamagishi A (2005) First-principles studies on the elastic constants of a 1:1 layered kaolinite mineral. *American Mineralogist* 90: 1824–1826.

Sayers CM (1999) Stress-dependent seismic anisotropy of shales. *Geophysics* 64: 93–98.

Sayers CM (2002) Stress-dependent elastic anisotropy of sandstones. *Geophysical Prospecting* 50: 85–95.

Sayers C (2004) Monitoring production – Induced stress changes using seismic waves. *SEG Technical Program Expanded Abstracts* 2004: 2287–2290.

Sayers CM (2007) Effects of borehole stress concentration on elastic wave velocities in sandstones. *International Journal of Rock Mechanics and Mining Sciences* 44: 1045–1052.

Sayers CM and Den Boer LD (2011) Rock physics-based relations for density and *S*-velocity versus *P*-velocity in deepwater subsalt Gulf of Mexico shales. *The Leading Edge* 30: 1376–1381.

Sayers CM and Kachanov M (1995) Microcrack-induced elastic wave anisotropy of brittle rocks. *Journal of Geophysical Research, Solid Earth* 100: 4149–4156.

Sayers CM, Vanmunster JG, and King MS (1990) Stress-induced ultrasonic anisotropy in Berea sandstone. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* 27: 429–436.

Schijns H, Schmitt DR, Heikkinen PJ, and Kukkonen IT (2012) Seismic anisotropy in the crystalline upper crust: Observations and modelling from the outokumpu scientific borehole, finland. *Geophysical Journal International* 189: 541–553.

Schmitt DR (1999) Seismic attributes for monitoring of a shallow heated heavy oil reservoir: A case study. *Geophysics* 64: 368–377.

Schmitt DR, Han Z, Kravchinsky VA, and Escartin J (2007) Seismic and magnetic anisotropy of serpentinized ophiolite: Implications for oceanic spreading rate dependent anisotropy. *Earth and Planetary Science Letters* 261: 590–601.

Schmitt DR and Li Y (1995) A high-pressure technique for determining the microcrack porosities of damaged brittle materials. *Canadian Journal of Physics* 73: 330–337.

Schmitt DR, Mwenifumbo J, Pflug KA, and Meglis IL (2003) Geophysical logging for elastic properties in hard rock: A tutorial. In: Eaton DW, Milkereit B, and Salisbury MH (eds.) *Hardrock Seismic Exploration*, pp. 20–41. Tulsa: Society of Exploration Geophysicists.

Schmitt DR, Smither C, and Ahrens TJ (1989) In situ holographic elastic-moduli measurements from boreholes. *Geophysics* 54: 468–477.

Schmitt DR and Zoback MD (1992) Diminished pore pressure in low-porosity crystalline rock under tensional failure – Apparent strengthening by dilatancy. *Journal of Geophysical Research, Solid Earth* 97: 273–288.

Schubnel A, Benson PM, Thompson BD, Hazzard JF, and Young RP (2006) Quantifying damage, saturation and anisotropy in cracked rocks by inverting elastic wave velocities. *Pure and Applied Geophysics* 163: 947–973.

Schubnel A and Gueguen Y (2003) Dispersion and anisotropy of elastic waves in cracked rocks. *Journal of Geophysical Research, Solid Earth* 108: 2101–2116.

Setzmann U and Wagner W (1989) A new method for optimizing the structure of thermodynamic correlation equations. *International Journal of Thermophysics* 10: 1103–1126.

Setzmann U and Wagner W (1991) A new equation of state and tables of thermodynamic properties for methane covering the range from the melting line to 625-K at pressures up to 1000-MPa. *Journal of Physical and Chemical Reference Data* 20: 1061–1155.

Sharqawy MH, Lienhard JH, and Zubair SM (2010) Thermophysical properties of seawater: A review of existing correlations and data. *Desalination and Water Treatment* 16: 354–380.

Simmons G and Wang HF (1971) *Single Crystal Elastic Constants and Calculated Aggregate Properties: A Handbook*. Cambridge, MA: MIT Press.

Sinha BK and Kostek S (1996) Stress-induced azimuthal anisotropy in borehole flexural waves. *Geophysics* 61: 1899–1907.

Smeulders DMJ (2005) Experimental evidence for slow compressional waves. *Journal of Engineering Mechanics, ASCE* 131: 908–917.

Smith TM, Sondergeld CH, and Rai CS (2003) Gassmann fluid substitutions: A tutorial. *Geophysics* 68: 430–440.

Smith T, Sondergeld C, and Tinni A (2010) Microstructural controls on electric and acoustic properties in tight gas sandstones; some empirical data and observations. *The Leading Edge* 29: 1470–1474.

Smyth JR and Mccormick TC (1995) Crystallographic data for minerals. In: Ahrens TJ (ed.) *Mineral Physics and Crystallography: A Handbook of Physical Constants*, pp. 1–17. Washington, DC: American Geophysical Union.

Smyth JR and Mccormick TC (2013) Crystallographic data for minerals. In: *Mineral Physics & Crystallography: A Handbook of Physical Constants*, pp. 1–17. Washington, DC: American Geophysical Union.

Span R and Wagner W (1996) A new equation of state for carbon dioxide covering the fluid region from the triple-point temperature to 1100 K at pressures up to 800 MPa. *Journal of Physical and Chemical Reference Data* 25: 1509–1596.

Spycher N, Pruess K, and Ennis-King J (2003) CO_2–H_2O mixtures in the geological sequestration of CO_2. I. Assessment and calculation of mutual solubilities from 12 to 100 degrees c and up to 600 bar. *Geochimica et Cosmochimica Acta* 67: 3015–3031.

Stanek M, Geraud Y, Lexa O, Spacek P, Ulrich S, and Diraison M (2013) Elastic anisotropy and pore space geometry of schlieren granite: Direct 3-D measurements at high confining pressure combined with microfabric analysis. *Geophysical Journal International* 194: 383–394.

Stein S and Wysession M (2002) *An Introduction to Seismology, Earthquakes, and Earth Structure*. New York: Wiley-Blackwell.

Stroisz AM and Fjaer E (2013) Tracing causes for the stress sensitivity of elastic wave velocities in dry castlegate sandstone. *Geophysical Journal International* 192: 137–147.

Sumino Y and Anderson OL (1984) Elastic constants of minerals. In: Carmichael RS (ed.) *Handbook of Physical Properties or Rocks, III* Boca Raton, FL: CRC Press.

Tadmor EB, Miller RE, and Elliot RS (2012) *Continuum Mechanics and Thermodynamics: From Fundamental Concepts to Governing Equations*. New York: Cambridge University Press.

Takenouchi S and Kennedy GC (1964) Binary system H_2O–CO_2 at high temperatures +pressures. *American Journal of Science* 262: 1055–1074.

Thomas LK, Hankinson RW, and Phillips KA (1970) Determination of acoustic velocities for natural gas. *Journal of Petroleum Technology* 22: 889–892.

Thomsen L (1972) Elasticity of polycrystals and rocks. *Journal of Geophysical Research* 77: 315–327.

Tinder RF (2007) Tensor properties of solids, part two: Transport properties of solids. *Synthesis Lectures on Engineering*, vol. 2, pp. 145–236.

Todt T and Simmons G (1972) Effect of pore pressure on velocity of compressional waves in low-porosity rocks. *Journal of Geophysical Research* 77: 3731–3743.
Umekawa S and Sherby O (1966) Young's modulus of eutectic alloy systems with special emphasis on silver-lead system. *Journal of the Mechanics and Physics of Solids* 14: 65.
van Der Waals JD (1873) *Over de continuiteit van den gas- en vloeistoftoestand (on the continuity of the gas and liquid state)*. PhD, Leiden.
Van Weert F and Van Der Gun J (2012) Saline and brackish groundwater at shallow and intermediate depths: Genesis and worldwide occurrence. In: *39th International Association of Hydrologists*, Niagara Falls: International Association of Hydrologists.
Vance S and Brown JM (2010) Sound velocities and thermodynamic properties of water to 700 MPa and −10 to 100 degrees C. *Journal of the Acoustical Society of America* 127: 174–180.
Vankonynenburg PH and Scott RL (1980) Critical lines and phase equilibria in binary van der waals mixtures. *Philosophical Transactions of the Royal Society Series A: Mathematical Physical and Engineering Sciences* 298: 495–540.
Voigt W (1887) Theoretische studien über die elasticitätsverhältnisse der krystalle. *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen* 34: 3–51.
Voigt W (1928) *Lehrbuch der kristallphysik (mit ausschluss der kristalloptik)*. Leipzig: B.G. Teubner.
Wagner W (2013) *Software for the Reference Equation of State Gerg-2008 for Natural Gases and Other Mixtures [Online]*. Bochum: Ruhr-Universität Bochum, Available: http://www.thermo.rub.de/en/prof-w-wagner/software/gerg-2004-gerg-2008.html.
Wagner W and Pruss A (2002) The iapws formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. *Journal of Physical and Chemical Reference Data* 31: 387–535.
Walker AM and Wookey J (2012) Msat-a new toolkit for the analysis of elastic and seismic anisotropy. *Computers & Geosciences* 49: 81–90.
Walsh JB (1965) Effect of cracks on compressibility of rock. *Journal of Geophysical Research* 70: 381–389.
Walsh JB, Brace WF, and England AW (1965) Effect of porosity on compressibility of glass. *Journal of the American Ceramic Society* 48: 605–608.
Walton K (1987) The effective elastic-moduli of a random packing of spheres. *Journal of the Mechanics and Physics of Solids* 35: 213–226.
Wang HF (2000) *The Linear Theory of Poroelasticity with Applications to Geomechanics and Hydrology*. Princeton: Princeton University Press.
Wang ZJ and Nur A (1991) Ultrasonic velocities in pure hydrocarbons and mixtures. *Journal of the Acoustical Society of America* 89: 2725–2730.
Wang ZJ, Nur AM, and Batzle ML (1990) Acoustic velocities in petroleum oils. *Journal of Petroleum Technology* 42: 192–200.
Watt JP (1979) Hashin–shtrikman bounds on the effective elastic-moduli of polycrystals with orthorhombic symmetry. *Journal of Applied Physics* 50: 6290–6295.
Watt JP (1980) Hashin–shtrikman bounds on the effective elastic-moduli of polycrystals with monoclinic symmetry. *Journal of Applied Physics* 51: 1520–1524.
Watt JP (1986) Hashin–shtrikman bounds on the effective elastic-moduli of polycrystals with trigonal (3,3 bar) and tetragonal (4,4 bar,4m) symmetry. *Journal of Applied Physics* 60: 3120–3124.
Watt JP, Davies GF, and O'Connell RJ (1976) Elastic properties of composite-materials. *Reviews of Geophysics* 14: 541–563.
Watt JP and Peselnick L (1980) Clarification of the Hashin–Shtrikman bounds on the effective elastic-moduli of polycrystals with hexagonal, trigonal, and tetragonal symmetries. *Journal of Applied Physics* 51: 1525–1531.
Wendland M, Hasse H, and Maurer G (1999) Experimental pressure–temperature data on three- and four-phase equilibria of fluid, hydrate, and ice phases in the system carbon dioxide-water. *Journal of Chemical and Engineering Data* 44: 901–906.
Wenk HR, Lonardelli I, Franz H, Nihei K, and Nakagawa S (2007) Preferred orientation and elastic anisotropy of illite-rich shale. *Geophysics* 72: E69–E75.
Wenk HR, Vasin RN, Kern H, Matthies S, Vogel SC, and Ivankina TI (2012) Revisiting elastic anisotropy of biotite gneiss from the Outokumpu scientific drill hole based on new texture measurements and texture-based velocity calculations. *Tectonophysics* 570: 123–134.
White D (2013) Seismic characterization and time-lapse imaging during seven years of CO_2 flood in the Weyburn field, Saskatchewan, Canada. *International Journal of Greenhouse Gas Control* 16(supplement 1): S78–S94.
Wilson WD (1959) Speed of sound in distilled water as a function of temperature and pressure. *Journal of the Acoustical Society of America* 31: 1067–1072.
Wilson PS (2005) Low-frequency dispersion in bubbly liquids. *Acoustics Research Letters Online (ARLO)* 6: 188–194.
Winkler KW (1996) Azimuthal velocity variations caused by borehole stress concentrations. *Journal of Geophysical Research, Solid Earth* 101: 8615–8621.
Wohlenberg J (1982) Density of rocks. In: Hellwege K-H (ed.) *Landolt–Bornstein Numerical Data and Functional Relationships in Science and Technology*, pp. 11–199. Berlin: Springer-Verlag.
Wong GSK and Embleton TFW (1984) Variation of specific-heats and of specific-heat ratio in air with humidity. *Journal of the Acoustical Society of America* 76: 555–559.
Wong GSK and Zhu SM (1995) Speed of sound in seawater as a function of salinity, temperature, and pressure. *Journal of the Acoustical Society of America* 97: 1732–1736.
Wood AB (1955) *A Textbook of Sound*. London: George Bell & Sons.
Wu T (1966) The effect of inclusion shape on the elastic moduli of a two-phase material. *International Journal of Solids and Structures* 2: 1–8.
Wuestefeld A, Verdon JP, Kendall JM, Rutledge J, Clarke H, and Wookey J (2011) Inferring rock fracture evolution during reservoir stimulation from seismic anisotropy. *Geophysics* 76: WC157–WC166.
Wyllie M, Gregory A, and Gardner G (1958) An experimental investigation of factors affecting elastic wave velocities in porous media. *Geophysics* 23: 459–493.
Xu XX, Hofmann R, Batzle M, and Tshering T (2006) Influence of pore pressure on velocity in low-porosity sandstone: Implications for time-lapse feasibility and pore-pressure study. *Geophysical Prospecting* 54: 565–573.
Yam H (2011) *CO_2 Rock Physics: A Laboratory Study*. MSc, University of Alberta.
Yam H and Schmitt DR (2011) CO_2 rock physics: A laboratory study. *InSite Magazine* 30: 13–16.
Yan XF, Yao FC, Cao H, Ba J, Hu LL, and Yang ZF (2011) Analyzing the mid-low porosity sandstone dry frame in central Sichuan based on effective medium theory. *Applied Geophysics* 8: 163–170.
Yoshimura HN, Molisani AL, Narita NE, Cesar PF, and Goldstein H (2007) Porosity dependence of elastic constants in aluminum nitride ceramics. *Materials Research* 10: 127–133.
Ziabakhsh-Ganji Z and Kooi H (2012) An equation of state for thermodynamic equilibrium of gas mixtures and brines to allow simulation of the effects of impurities in subsurface CO_2 storage. *International Journal of Greenhouse Gas Control* 11: S21–S34.
Zimmer M (2003) *Seismic Velocities in Unconsolidated Sands: Measurements of Pressure, Sorting, and Compaction Effects*. PhD, Stanford University.
Zimmer M, Prasad M, Mavko G, and Nur A (2007) Seismic velocities of unconsolidated sands: Part 1 – Pressure trends from. *Geophysics* 72: E1–E13.
Zimmerman RW (1984) Elastic-moduli of a solid with spherical pores – New self-consistent method. *International Journal of Rock Mechanics and Mining Sciences* 21: 339–343.
Zimmerman RW (1985) The effect of microcracks on the elastic-moduli of brittle materials. *Journal of Materials Science Letters* 4: 1457–1460, *AAPG Bulletin-American Association of Petroleum Geologists* 59: 154–158.
Zoback MD and Byerlee JD (1975) Permeability and effective stress. *AAPG Bulletin-American Association of Petroleum Geologists* 59(1): 154–158.

This page intentionally left blank

11.04 Geophysical Properties of the Near Surface Earth: Electrical Properties

PWJ Glover, University of Leeds, Leeds, UK

© 2015 Elsevier B.V. All rights reserved.

11.04.1	**Introduction to the Electrical Properties of Near-Surface Rocks**	90
11.04.1.1	Importance	90
11.04.1.2	Scope	90
11.04.1.3	Objectives	91
11.04.1.4	Structure of the Chapter	91
11.04.2	**Steady-State Electrical Properties**	91
11.04.2.1	Basic Physical Relationships	91
11.04.2.1.1	Ohm's law	91
11.04.2.1.2	Ohm's law in a continuous medium	92
11.04.2.1.3	Resistance and resistivity	92
11.04.2.1.4	Conductance and conductivity	93
11.04.2.1.5	Conductivity and charge carriers	93
11.04.2.1.6	Electronic conduction	93
11.04.2.1.7	Ionic conduction	94
11.04.2.1.8	Conductivity of pore fluids	94
11.04.2.2	Parameters Affecting the Electrical Conductivity of Rocks	96
11.04.2.3	Archie's Original Empirical Laws	97
11.04.2.3.1	Formation factor and connectedness	98
11.04.2.3.2	Archie's first law	98
11.04.2.3.3	Resistivity index	100
11.04.2.3.4	Archie's second law	100
11.04.2.3.5	Combined Archie's laws	101
11.04.2.3.6	Electrical tortuosity and connectivity	101
11.04.2.3.7	Validity of Archie's laws	101
11.04.2.4	Empirical Modifications for Shaley Rocks	102
11.04.2.5	The Range of Electrical Conductivity of Near-Surface Rocks	102
11.04.3	**Mixing Models for the Electrical Conductivity of Geomaterials**	103
11.04.3.1	Geometric Models	103
11.04.3.2	Models with Variable Exponents	105
11.04.3.3	Extended and Generalized Archie's Laws	106
11.04.3.4	Empirical Models for Full Water Saturation	107
11.04.3.5	Shaley Sand Water Saturation Equations	107
11.04.4	**Surface Conduction**	109
11.04.4.1	The Electrical Double Layer and Surface Conduction	109
11.04.4.1.1	The electrical double layer (EDL)	109
11.04.4.1.2	Thickness of the electrical double layer	110
11.04.4.2	Basic Theory of Surface Conduction	111
11.04.4.2.1	Waxman and Smits model	111
11.04.4.2.2	Ionic surface electrical conduction models	111
11.04.4.3	Application to Clean Rocks	111
11.04.4.4	Application to Clay-Rich Rocks	112
11.04.5	**Frequency-Dependent Electrical Properties**	112
11.04.5.1	Complex Conduction	113
11.04.5.2	Dielectric Permittivity and Polarization	113
11.04.5.3	Dielectric Permittivity Models	114
11.04.5.4	Integrated Conduction and Permittivity	116
11.04.5.5	Mixing Models for Complex Electrical Properties	117
11.04.5.5.1	The CRIM model	117
11.04.5.5.2	The Lichtenecker–Rother equation	117
11.04.5.6	Modeling the Frequency-Dependent Electrical Properties of Rocks	117
11.04.5.7	Measurement of Frequency-Dependent Electrical Properties	119
11.04.5.8	Dielectric Properties of Pore Fluids	119
11.04.5.9	Dielectric Properties of Minerals and Near-Surface Rocks	120
11.04.5.9.1	Minerals	120

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00189-5

11.04.5.9.2	Rocks	120
11.04.6	**Electrokinetic Properties**	121
11.04.6.1	Streaming Potentials	121
11.04.6.2	Electroosmosis	122
11.04.6.3	Seismoelectrics	122
11.04.6.4	The Helmholtz–Smoluchowski Equation in the Steady State	122
11.04.6.5	Calculating Zeta Potentials from Experimental Measurements	123
11.04.6.6	Grain, Pore, and Pore Throat Sizes and Streaming Potential	123
11.04.6.7	Streaming Potential and Zeta Potential Models	123
11.04.6.8	Factors Affecting the Streaming Potential of Near-Surface Rocks	125
11.04.6.9	Electrokinetic Measurements under Steady-State Conditions	126
11.04.6.9.1	Experimental considerations	126
11.04.6.9.2	Salinity and pore fluid conductivity	126
11.04.6.9.3	Pore fluid pH	126
11.04.6.9.4	Pore fluid composition	127
11.04.6.9.5	Mineralogy	127
11.04.6.9.6	Rock microstructure	128
11.04.6.9.7	Saturation	128
11.04.6.9.8	Temperature	129
11.04.6.10	Streaming Potential as a Function of Frequency	129
11.04.7	**Summary**	131
11.04.7.1	History and Progress	131
11.04.7.2	Current Challenges and Future Directions	131
References		132

11.04.1 Introduction to the Electrical Properties of Near-Surface Rocks

11.04.1.1 Importance

The electrical properties of near-surface geomaterials have been hugely important in the development of many modern well-developed countries, and will play a leading role in the further development of those less well-developed countries which are currently emerging as possible future economic powers. The effect is, of course, most obvious in the exploitation of hydrocarbon resources, where Archie's laws have been at the heart of oil and gas reserves calculations for every field that has been produced since their inception (Archie, 1942). This amounts to more than 1.5 trillion barrels of oil and 7.5 trillion cubic feet of gas (Bentley, 2002). These resources have driven the global economy of the last 50 years (Glover, 2009; Speight, 2011). The impact of oil and gas is overwhelming. According to the International Energy Agency's Energy Balance data, hydrocarbons heat the majority of all we heat, transport most of that which is transported, and make a significant proportion of our manufactured goods. Modern health, for example, would be impossible without the sterile plastic consumables upon which it depends.

However, it would be a mistake to consider the role of electrical conduction in near-surface materials to be confined to the oil and gas industry. There are significant applications in mining (e.g., Friedrichs et al., 1999); soil remediation (e.g., Pearce and Zuluaga, 2004; Rojo et al., 2014); waste management (domestic, industrial, toxic, and radioactive) (e.g., Lesparre et al., 2013); civil engineering, where the artificial rock concrete is a major factor in the development of our modern landscape (e.g., Susanto et al., 2013; Weiss et al., 2013); treatment of natural hazards including rock and mud slides (e.g., Jackson et al., 2002); and earthquakes and volcanic activity (e.g., Di Maio and Patella, 1991, 1994; Fujinawa et al., 1992; Mizutani et al., 1976; Revil et al., 2002, 2010).

In a more academic sphere, the applications of the concepts described in this chapter are used for understanding the continental crust (e.g., Glover, 1996; Glover and Adám, 2008; Glover and Vine, 1992, 1995; Glover et al., 2000b; Yang, 2011) and oceanic crust (e.g., Baba, 2005; Gung et al., 2003; Key et al., 2013; Yoshino et al., 2008); the triggering of natural and synthetic earthquakes (e.g., Chelidze et al., 2003; Cyr et al., 2010); volcanic activity monitoring (e.g., Revil et al., 2010); the properties of lunar (e.g., Li et al., 2005), Martian (e.g., Carter et al., 2009), and asteroidal (e.g., Wittmann et al., 1999) materials; and soil science (e.g., Robinson et al., 2003; Peplinski et al., 1995). Counterintuitively, perhaps, some of the electrical conductivity models have been applied in fields as varied as catalysis research (e.g., Zhang and Catchmark, 2011) and food engineering (e.g., Jha et al., 2011; Jindal et al., 2013). Certain concepts such as cation-exchange capacities (CECs) even find themselves useful in farming and gardening (e.g., Banton et al., 1997; Green, 1997; Oliver et al., 2013). Electrical measurements are also used by archaeometrists to search for buried archaeological remains (e.g., Gaffney, 2008; Glover, 2010a), as well as in the study and conservation of items of cultural heritage (e.g., Maurício et al., 2005; Sass and Viles, 2010; Zhao et al., 2012).

11.04.1.2 Scope

The electrical properties of rocks describe how charge is transported through them. The transport of charges may be in response to a steady-state or time-varying electric field, giving rise to a steady-state resistivity and conductivity or a complex

impedance or admittance. In each case, the charge carriers may be electrons or ions.

Electrons are the charge carriers that mediate conduction in metals and semiconductors. It is the conduction mechanism when conduction takes place in minerals at high temperatures and pressures and falls outside the scope of this volume (please see Chapter 2.25).

Ions are the charge carriers that mediate conduction in pore fluids, in the electrical double layer that exists at the mineral/pore fluid interface, and in melts. All three situations are common in near-surface rocks and are considered in detail in this chapter. In the steady state, we consider the electrical resistivity and conductivity, but use complex impedance and admittance when discussing the frequency-dependent conduction, each of which is composed of an in-phase resistivity or conductivity and an out-of-phase or quadrature resistivity or conductivity.

There may be net transport of ionic charges in the absence of an electric field. For example, preferential advection of positive and negative ions may take place due to the flow of fluids. The preferential advection then causes a separation of charge that results in a current source. The streaming potential is the electric field associated with that current source. This creation of an electric potential in response to a fluid flow in the absence of an external applied electric field is an example of an electrokinetic process. Electrokinetic processes are described towards the end of this chapter.

11.04.1.3 Objectives

This chapter has the following objectives:

- To describe the fundamental physics that underlies the steady-state conduction of electric currents in geomaterials including conduction mechanisms.
- To examine the electrical properties of the various materials that contribute to the electrical properties of geomaterials.
- To review briefly the range of resistivity and conductivity values found in common geomaterials.
- To introduce Archie's laws in detail, noting their important and wide-ranging academic and industrial applications in understanding near-surface geomaterials.
- To examine theoretical and empirical models for steady-state conduction in geomaterials including those for partially saturated and shaley rocks.
- To examine the origins, role, and implications of surface conduction in geomaterials and to review briefly some applications for clean and shaley geomaterials.
- To describe the measurement, modeling, and range of frequency-dependent conduction in geomaterials, including a review of the dispersion processes that control the overall dielectric properties of the geomaterial, measurement, and modeling, together with the range of dielectric properties of geomaterials and the materials that compose them.
- To describe the electrokinetic processes that couple the electrical properties of rocks to fluid flowing through them and the main parameters that affect the electrokinetic coupling.
- To describe the measurement and modeling of these electrokinetic properties both in the steady state and as a function of frequency.

11.04.1.4 Structure of the Chapter

This chapter is in six parts. Section 11.04.1 serves as a short general introduction. Section 11.04.2 describes steady-state electrical properties, starting with Ohm's law, discussing the properties that affect the electrical properties of rocks, introducing Archie's laws, and examining the range of electrical properties encountered in rocks and minerals in the natural world.

There are many mixing models that aim to provide the conductivity or resistivity of a geomaterial once the conductivity or resistivity of its components and their volume fractions are known. Section 11.04.3 examines mixing models in detail, making a distinction between geometric arrangements of phases and those that invoke parameters to describe the connectedness of each phase.

Section 11.04.4 examines surface conduction in detail, introducing the electrical double layer and describing the basic theory of surface conduction and its applications to clean and clay-containing geomaterials.

Steady-state conduction is a special case of conduction where the electric potential and current remain constant with time. In general, this is not so. Electromagnetic fields may vary harmonically at a given frequency or transiently. Section 11.04.5 describes frequency-dependent conduction in geomaterials, starting with an understanding of the dielectric properties of geomaterials and how they provoke polarization mechanisms within geomaterials and how conduction and displacement currents contribute to the overall complex electrical properties of the geomaterial. Mixing models, theoretical modeling, and measurement (impedance spectroscopy) of the complex electrical properties of geomaterials are also considered before examining the range of dielectric properties of geofluids and geomaterials.

If a porous medium has an electrical double layer at the fluid–matrix interface, it develops a finite surface conduction. When fluid flows in such a medium, there is an electrokinetic coupling that generates an electrical streaming potential. Section 11.04.6 examines the origin, properties, measurement, and modeling of the electrokinetic properties of geomaterials.

Finally, Section 11.04.7 provides a brief summary.

11.04.2 Steady-State Electrical Properties

11.04.2.1 Basic Physical Relationships

11.04.2.1.1 Ohm's law

Ohm's law is the fundamental physical law that relates the current flowing through a material to the difference in potential across it.

The law states that in an electrical circuit, the current (I, in A) passing through a conductor between two points is directly proportional to the difference in electric potential (V, in V) between the two points and inversely proportional to the resistance of the material (R, in Ω) between the two points:

$$I = \frac{V}{R} \quad [1]$$

The potential difference represents the electric force that causes a flow of charge carriers that forms the current. The resistance is that property of the material which resists the flow of the charge carriers, and limits current density proportionally to the electric potential difference.

11.04.2.1.2 Ohm's law in a continuous medium

At any point in a continuous medium, Ohm's law can be written as

$$\vec{J} = \sigma \vec{E} \quad [2]$$

where $\vec{J}$ is the current density, in A m^{-2}; σ is the conductivity of the material, in S m^{-1}; and $\vec{E}$ is the electric field, in V m^{-1}.

The conductivity, σ, is usually a scalar function that may vary with angular frequency, but can be a nonlinear tensor function of temperature, pressure, and deformation as well as being a tensor due to anisotropy.

If the material moves with a velocity $\vec{v}$, relative to a magnetic field $\vec{B}$, the current density is augmented by a contribution that is induced in the material by the magnetic field:

$$\vec{J} = \sigma\left(\vec{E} + \vec{v} \times \vec{B}\right) \quad [3]$$

This relationship is not developed in the remainder of this chapter, but is an important starting point for understanding geophysical techniques such as magnetotellurics.

11.04.2.1.3 Resistance and resistivity

The resistance (R) is the ability of a material to impede the passage of an electric current. It is measured in ohms (Ω). The resistance depends upon the size and shape of the material in which the current flows. Doubling the distance through which the current flows doubles the resistance, while doubling the cross-sectional area of material perpendicular to flow halves the resistance. Hence, resistance is partly due to some intrinsic property of the material and partly due to the geometry of the material relative to the two points between which the current flows and the electric potential is applied.

The resistivity (ρ) describes only the intrinsic resistive properties of the material to the flow of charge carriers and is not affected by the geometry of the sample. It corresponds to the resistance of a sample 1 m long with a constant cross-sectional area for current flow that is 1 m^2. Resistivity has units of ohm meters (Ωm).

For a continuous current that passes through a homogeneous and isotropic material with a length L and a cross-sectional area A, the resistance and resistivity are related by

$$\rho = R\frac{A}{L} \quad [4]$$

We can define the resistivity of a material as

$$\rho = \frac{\vec{E}}{\vec{J}} \quad [5]$$

where $\vec{E}$ is the magnitude of the electric field applied to the material (in V m^{-1}) and $\vec{J}$ is the magnitude of the current density that results (in A m^{-2}).

Resistivity depends on temperature. As a general rule, the resistivity of a metal increases with temperature, while that of a semiconductor decreases with temperature. The resistance of minerals at high temperatures is beyond the scope of this chapter (see Chapter 2.25). The resistivity of aqueous fluids, which occupy the pores of near-surface rocks and form one of the main conduction pathways in such rocks, decreases with increasing temperature at about 4% per °C at low temperatures (around 5 °C), at about 2% per °C at room temperatures, and at about 1% per °C as the temperature approaches 100 °C (Sen and Goode, 1992a,b). This is a significant sensitivity to temperature and implies that all electrical measurements should be made at a stable temperature, and care is taken comparing measurements made at different temperatures. Conversely, some field applications use resistivity to infer the in situ temperature, for example, some geothermal applications and the monitoring of groundwater-surface water exchange.

Figure 1 shows the electrical resistivity of several types of Earth material. It is worth noting that the range of measured resistivities is extremely large from $10^{-7}\Omega$m for copper to $10^{18}\Omega$m for diamond (Pierson, 2004). However, electrical resistivity measurements may be made in the laboratory with a precision of several decimal places. Hence, electrical

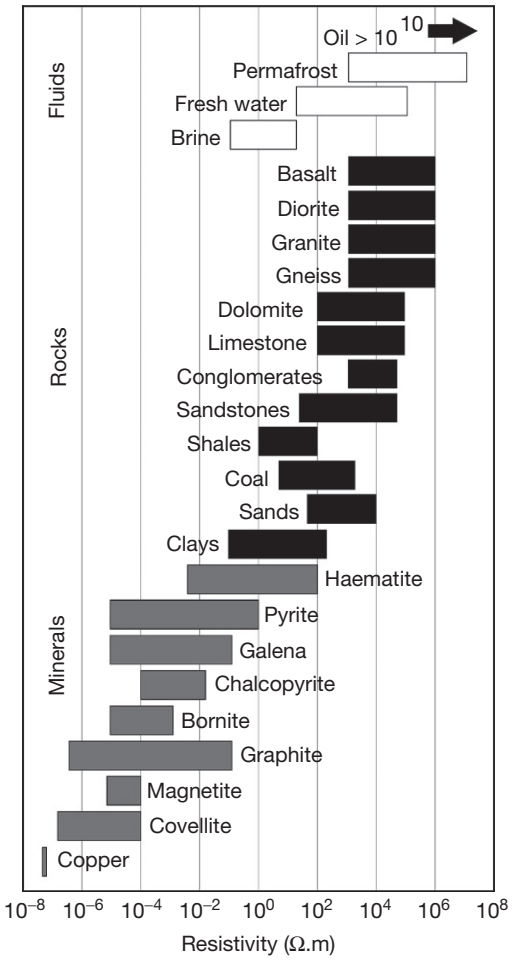


Figure 1 Electrical resistivity of several geomaterials (compiled from Guéguen and Palciauskas, 1994; Schön, 2004; Keller, 1989).

properties offer a very precise way of differentiating geomaterials and can be used to develop probes, which are highly sensitive to changes in rock structure and composition. However, many geomaterials have a large natural range of electrical properties that arise from their natural heterogeneity and anisotropy. When this is the case, even highly sensitive electrical techniques cannot differentiate between them.

Electrical measurements have the additional advantage in that they can often be made remotely. A small-scale example is the electrical induction downhole tool, which can be used to measure the conductivity of a formation even though the well is filled with an insulating oil-based mud. A large-scale example would be the magnetotelluric method that allows the conductance of layers deep in the Earth to be measured. In both cases, a current is induced in the material that is to be measured using an electromagnetic field. The current that flows induces a secondary field, which itself can be measured in order to obtain the resistance of the remote material.

11.04.2.1.4 Conductance and conductivity

Conductance (G_c), which is expressed in siemens (S), is the inverse of resistance (R). Hence (Guéguen and Palciauskas, 1994),

$$G_c = \frac{1}{R} \quad [6]$$

Conductance represents the ability of a material to let charges flow through it freely.

Conductivity (σ) is, therefore, the inverse of resistivity:

$$\sigma = \frac{\vec{J}}{\vec{E}} \quad [7]$$

where $\vec{J}$ is the current density that flows in the material, in A m^{-2}, and $\vec{E}$ is the electric field, in V m^{-1}. Conductivity is expressed in siemens per meter (S m^{-1}).

11.04.2.1.5 Conductivity and charge carriers

Consider charge carrier i with a charge $Z_i e$, where Z_i is the valency of the charge carrier and $e \approx 1.6022 \times 10^{-19}$ C is the elementary charge (Lide, 2012). If the charge carrier is in an electric field $\vec{E}$, the force on the charge carrier is $Z_i e \vec{E}$. In a steady state, the force results in a uniform charge carrier velocity:

$$\vec{v} = \mu_i \vec{E} \quad [8]$$

where μ_i is the mobility of the charge carrier, in m^2 V^{-1} s^{-1}, and is dependent upon its type, the medium, and temperature. Since the electric current density is the charge that crosses a unit surface area per second and n_i is the number of charge carriers of type i per unit volume,

$$\vec{J} = n_i Z_i e \vec{v} \quad [9]$$

Combining [7] to [9] gives an expression for conductivity as

$$\sigma_i = n_i Z_i e \mu_i \quad [10]$$

The conductivity is directly proportional to (i) the number of charge carriers per unit volume, (ii) the charge that each carries, and (iii) the mobility of the charge carriers. This basic statement is true for electronic conduction in solids; ionic conduction in solids, aqueous fluids, and melts; and for ionic conduction within the electrical double layer at the mineral/fluid interface.

11.04.2.1.6 Electronic conduction

All materials contain electrons. However, not all electrons are mobile. In solids, electrons have well-defined energy states that form bands. These bands are separated by disallowed energy states. Electrons may be mobile if they occupy a partially filled band of energy states. Then, externally applied energy, for example, from the ambient temperature ($=k_bT$) or an externally applied electric field ($= e\vec{E}$), can promote an electron to a higher energy, which is reflected in its kinetic energy (i.e., its motion).

Good conductors such as metals have a partially filled energy band, and the addition of externally applied energy results in the motion of electrons in this 'conduction' band (Poirier, 2000). By contrast, insulators have a full energy band so no externally supplied energy can be gained by electrons, and consequently, electrons cannot move. Moreover, the disallowed energy states that exist between the full band and the next, empty, conduction band is so large that electrons cannot be promoted to it and hence cannot gain kinetic energy of movement (Poirier, 2000).

For semiconductors, the energy band is also full. However, the energy gap between it and the next allowed energy band is small and electrons may be promoted into the higher-energy conduction band (Poirier, 2000). At a given temperature T, a number n_e of electrons can be promoted (excited thermally) into the conduction band according to

$$n_e \approx \exp\left\{-\frac{E_g}{k_b T}\right\} \quad [11]$$

where E_g is the width of the energy gap between the bands. If $E_g \gg k_b T$, there are very few electrons in the conduction band and the conductance is negligibly small. For minerals with $E_g \approx 0.5$ eV such as olivine ($E_g = 0.5$ eV; Guéguen and Palciauskas, 1994) or galena ($E_g = 0.37$ eV; Xu and Schoonen, 2000), there is only one promoted electron per 300 hundred million atoms of the mineral at 25 °C, but this rises to one promoted electron per 223 atoms at 800 °C, while a mineral with a large bandgap such as diamond ($E_g = 5.5$ eV; Guéguen and Palciauskas, 1994) or zirconia ($E_g = 5.0$ eV; Xu and Schoonen, 2000) has fewer than one conduction electron per 10^{12} atoms for all temperatures less than 2000 °C.

Hence, the electronic conductivity of a solid depends very much upon the size of the bandgap and the temperature, which control the number of electrons that are mobile and can take part in conduction according to eqn [10]. This equation shows that the conductivity also depends on the mobility of the electrons.

The electron mobility is given by

$$\mu_e = \frac{e\tau}{m_e} \quad [12]$$

where τ is the time between collisions between the electrons and the lattice and m_e is the mass of the electron. Collisions become more frequent with increasing temperature, and consequently, electron mobility diminishes. However, this effect is much smaller than the effect of temperature on the number of

electrons available for conduction, and hence, the overall trend is that the conductivity of solids increases strongly with temperature.

The conductivity of insulators and semiconductors also increases slightly with pressure due to a reduction in the band-gap energy. The effect of pressure is extremely small compared with that of temperature, and can be neglected in the Earth's crust (Dai and Karato, 2009; Xu et al., 2000). If one takes a typical geotherm of 30 °C km^{-1} and a typical geobar of 30 MPa km^{-1}, the effect of pressure is less than 1% of the effect of temperature at upper crustal conditions.

11.04.2.1.7 Ionic conduction

Transport of charge is mediated by ions. Ionic conduction is different from electronic conduction in that there are fewer ions available to carry charges than there are electrons in a metal or semiconductor, and since ions of different charge impede each other and even bond to each other, their mobilities are smaller.

Conduction in pore fluids is mediated by charged ions. In an aqueous solution of NaCl, for example, there are finite nonzero concentrations of the charged ions H^+, H_3O^+, OH^-, Na^+, and Cl^-, all of which transfer charge when an electric potential difference is applied to the fluid.

The general equation for describing the flow of charge carriers in a direction x (or indeed any particle, whether charged or not, and whether of ionic size or macroscopic size) is given by

$$\vec{J} = -D_i \frac{\partial n_i}{\partial x} + \bar{v} n_i \qquad [13]$$

where the first term is a statement of Fick's law, n_i is the number density of ionic species i, and D_i is the diffusion coefficient of the charged species, in m^2 s^{-1}, while the second term is due to fluid advection in direction x with a mean velocity $\bar{v}$. If we imagine a steady-state regime where $\vec{J} = 0$, hence,

$$D_i \frac{\partial n_i}{\partial x} = \bar{v} n_i \qquad [14]$$

and that is in thermodynamic equilibrium so that the spatial charge carrier distribution follows Boltzmann's law:

$$n(x) = n_o \exp\{-\varphi / k_b T\} \qquad [15]$$

where $k_b \approx 1.3806 \times 10^{-23}$ m^2 kg s^{-1} K^{-1} is Boltzmann's constant (Lide, 2012) and φ is either the chemical potential or the electric potential, leading ultimately to the equation

$$\frac{dn}{dx} = -\frac{n}{k_b T}\frac{d\varphi}{dx} = \frac{nF}{k_b T} \qquad [16]$$

where $F = -d\varphi/dx$ is the applied force responsible for the advective transport. Combination of [14] and [16] leads to the Nernst–Einstein equation (Poirier, 2000):

$$\frac{\bar{v}}{D} = \frac{F}{k_b T} \qquad [17]$$

Since the applied force for the advective transport $F = Z_i e \vec{E}$, we can combine [17] with [8] to obtain the ionic mobility as

$$\mu_i = \frac{Z_i e D_i}{k_b T} \qquad [18]$$

11.04.2.1.8 Conductivity of pore fluids

By combining [9] with [18], we may obtain the electrical conductivity for an ion in solution that is based on the Nernst–Einstein relationship:

$$\sigma_i = \frac{D_i Z_i^2 e^2 N n_i}{k_b T} \qquad [19]$$

where i represents the ith charge carrier; D_i is the diffusion coefficient of the charged species, in m^2 s^{-1}; Z_i is the valency of the charge carrier; $N \approx 6.022 \times 10^{23}$ mol^{-1} is Avogadro's number; n_i is the concentration of the charge carrier in mol l^{-1} (Lide, 2012); and T is the absolute temperature, in kelvin (K). It is worth noting that this equation shows an apparent inverse dependence on temperature, but at temperatures lower than about 250 °C, the conductivity of pore fluids increases with temperature because other terms in eqn [19] are also dependent on temperature, more than counteracting the explicit inverse dependence.

In a pore fluid, the current is carried by anions and cations traveling in opposite directions. In a steady state, the driving electric force $Z_i e \vec{E}$ is exactly balanced by the viscous drag on the hydrated ion $6\pi\eta r \vec{v}$ that is, according to Stokes' law, proportional to the velocity of the ion $\vec{v}_i$ (in m s^{-1}) and its hydrated radius r_i (in m) as well as the viscosity of the fluid η (in Pa s). Hence (Guéguen and Palciauskas, 1994),

$$Z_i e \vec{E} = 6\pi\eta r_i \vec{v}_i \qquad [20]$$

Comparison of this equation with [8] shows that the ionic mobility can be written as

$$\mu_i = \frac{Z_i e}{6\pi\eta r} \qquad [21]$$

Hence, eqn [21] shows that the individual ionic mobilities depend upon the viscosity of the fluid. This will have importance later in this chapter when we discuss how the increasing viscosity as one approaches the Stern plane will reduce the ionic mobilities in the electrical double layer that exists at the mineral/pore fluid boundary. It should be noted that while both eqns [18] and [21] are often applied to ionic solutions such as pore fluids, eqn [21] relies on the assumption that the hydrated ions behave as spherical particles moving slowly through a fluid, and this is not necessarily the case at high salinities.

Equation [21] can be used with eqn [10] to obtain the conductivity of the ion in the solution:

$$\sigma_i = \frac{n_i Z_i^2 e^2}{6\pi\eta r} \qquad [22]$$

This equation is important as it clearly shows that the conductivity is

- directly proportional to the number density of charge carriers available to transport charge,
- directly proportional to the square of the charge carried ($Z_i^2 e^2$),
- inversely proportional to the radius of the hydrated ion, and
- inversely proportional to the viscosity of the fluid.

Hence, eqn [22] lays the foundation for having different conductivity contributions from anions and cations of different hydrated radius.

While eqns [19]–[22] describe the conductivity of a pore fluid formally, in practice, it is often obtained using plots, nomograms, and empirical formulae. Perhaps, the most common and useful empirical formula for the conductivity of an aqueous solution of NaCl is given by Sen and Goode (1992a,b), noting that the second of the two references corrects an important error in the relationship presented by the first. The relationship is

$$\sigma_f(T, C_f) = \left(d_1 + d_2 T + d_3 T^2\right) C_f - \left(\frac{d_4 + d_5 T}{1 + d_6 \sqrt{C_f}}\right) C_f^{3/2} \qquad [23]$$

where $d_1 = 5.6$ (S l m^{-1} mol^{-1}), $d_2 = 0.27$ (S l m^{-1} mol^{-1})/°C, $d_3 = -1.51 \times 10^{-4}$ (S l m^{-1} mol^{-1})/°C^2, $d_4 = 2.36$ (S m^{-1}/(mol l^{-1})$^{3/2}$), $d_5 = 0.099$ (S m^{-1}/(mol l^{-1})$^{3/2}$/°C), $d_6 = 0.214$ ((mol l^{-1})$^{-1/2}$), T is in °C, and C_f is the salinity of the bulk pore fluid (mol l^{-1}).

Hilchie's equation (Hilchie, 1984) is another method for calculating the conductivity of an NaCl fluid, but in this case, the equation is capable of converting fluid conductivities between temperature and salinities rather than being a formula for giving the conductivity directly.

Figures 2 and 3 show the values of eqn [23] for a range of fluid salinities that are found in the Earth's crust. Both salinity and temperature have a significant effect. Figure 2 shows the overwhelming effect of the number density of ions on the pore fluid conductivity. Fluids that are more saline have a greater number of ions and hence a greater conductivity. However, at very high salinities, interaction between ions begins to reduce their mobility, which reduces the rate of increase of conductivity with salinity.

The variation with temperature (Figure 3) is more complex. At low temperatures ($T < 140$ °C), the viscosity decreases with temperature and that leads to an increase in conductivity with temperature. At high temperatures ($T > 400$ °C), which is outside the range covered in Figure 3, the viscosity increases with-temperature leading to a decrease of conductivity with temperature.

It should be noted that eqn [23] and Figures 2 and 3 are for aqueous solutions of NaCl only. Equation [22] indicates that different anions and cations will have different contributions to the overall conductivity of a solution because they have different hydrated radii.

The overall conductivity of a solution composed of a complex mixture of ions can be found by defining an equivalent

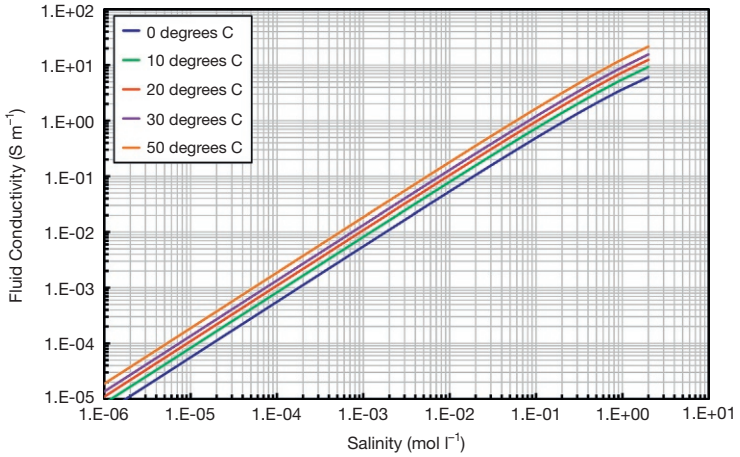


Figure 2 Conductivity of pore fluids as a function of salinity and temperature according to the empirical model of Sen and Goode (1992a,b).

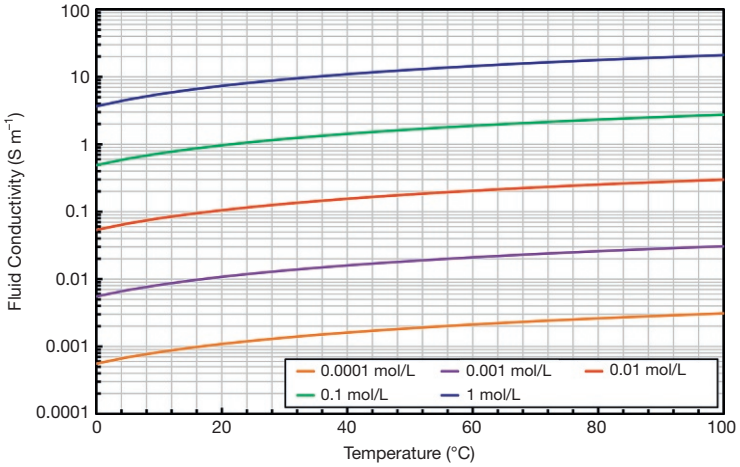


Figure 3 Conductivity of pore fluids as a function of temperature and salinity according to the empirical model of Sen and Goode (1992a,b).

NaCl solution that has the same conductivity as the complex solution. One of the most practical ways to do this is to define a multiplier for each of the ions in the solution except the Na^+ and Cl^- ions. These multipliers are called Dunlap coefficients. Because the salinity dependence of the conductivity of aqueous solutions of other ions is different from that of an NaCl solution, the multipliers also vary with salinity. Figure 4 shows the diagram that is used to choose these multipliers (Dunlap and Hawthorne, 1951; Schlumberger, 2013). It is best understood by taking an example: consider a solution that contains 20 000 ppm (parts per million) of NaCl, 10 000 ppm of KCl, and 1000 ppm of $MgSO_4$. The multipliers are Na(1.00), Cl(1.00), K(0.9), Mg(1.63), and SO_4(0.64), and the total NaCl equivalent solution is $(20\,000 \times 1 + 20\,000 \times 1 + 10\,000 \times 0.9 + 10\,000 \times 1 + 1000 \times 1.63 + 1000 \times 0.64)/2 = 30\,635$ ppm NaCl. Note that the sum of the products of the multipliers with the dissolved solids is divided by two because the calculation is carried out for both anions and cations. If there are dissolved salts that have two cations per anion or vice versa, the calculation becomes a little more complex. The value 30 635 ppm NaCl can now be entered into eqn [23] to obtain the conductivity of the complex pore fluid as a function of temperature.

Figure 4 covers the range 10 mg kg^{-1} to 300 000 mg kg^{-1}, which is approximately 0.0002–5 mol l^{-1}, taking NaCl as a typical reference.

In general, the use of Figure 4 followed by the application of the Sen and Goode equation is considered to be valid in this range of salinities. However, Figure 4 treats each ion as independent. In reality, different mixtures of ions interact to different extents and would lead to different effective NaCl concentrations. Furthermore, Figure 4 does not take into consideration that the multipliers will change with temperature. Hence, an effective NaCl concentration at the temperature for which Figure 4 is valid (25 °C) will be used in the Sen and Goode equation invoking a temperature of, say, 150 °C. An error inevitably results. Unfortunately, we only have one Figure 4 and it is for 25 °C.

One way of resolving the problem is to develop an electrochemical theory for the concentration dependence of ionic conductivity. This has been the subject of much study and the resulting Debye–Hückel theory is generally valid up to 10^{-2} mol l^{-1}.

Hence, if accurate values of conductivity are required up to 10^{-2} mol l^{-1}, the Debye–Hückel theory may be invoked, while for values in the range given by Figure 4, the empirical Sen and Goode method can be used providing that it is understood that its results are approximate.

Figure 5 shows the conductivity of pore fluids as a function of salinity for different pore fluid compositions derived from experimental measurements. It is clear that the general trend is sublinear with some electrolytes such as NaCl deviating from linear behavior at the highest salinities, while others such as KCl do not. It is worth noting that there is a variation of more than one order of magnitude in conductivity between the electrolytes included in this diagram at any given salinity. It is important, therefore, that the true composition of the pore fluid is taken into account in any laboratory or field measurement, especially if it changes spatially or temporally.

The small value of conductivity in bulk fluids at low salinities becomes important in geologic porous media because at these salinities, conduction through the electrical double layer becomes greater than that through the bulk fluid (see Section 11.04.2.4). The sensitivity to changes in temperature is also important because it implies that any experimental measurement of pore fluids or rocks saturated with pore fluids must be carried out at the same monitored temperature or be corrected to a standard temperature.

11.04.2.2 Parameters Affecting the Electrical Conductivity of Rocks

The electrical conductivity of rocks depends critically upon how current is carried through the rock. Deep in the Earth at high temperatures, conduction occurs through the solid minerals by electronic conduction, with minerals acting as semiconductors.

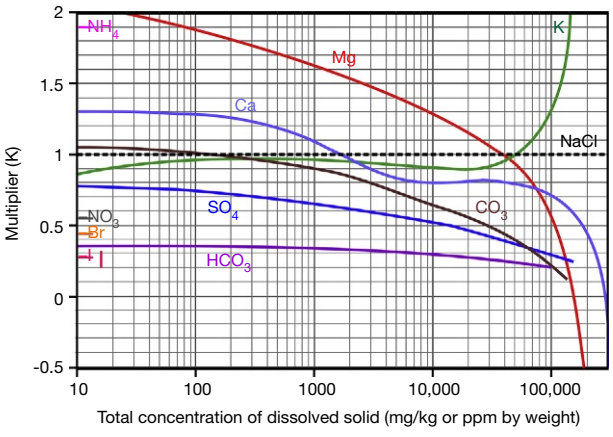


Figure 4 Graph of the multipliers used to convert a pore solution with a complex composition to an NaCl equivalent composition for the purposes of calculating the electrical conductivity of the solution. See the text, for example, of its use (Dunlap and Hawthorne, 1951; Schlumberger, 2013).

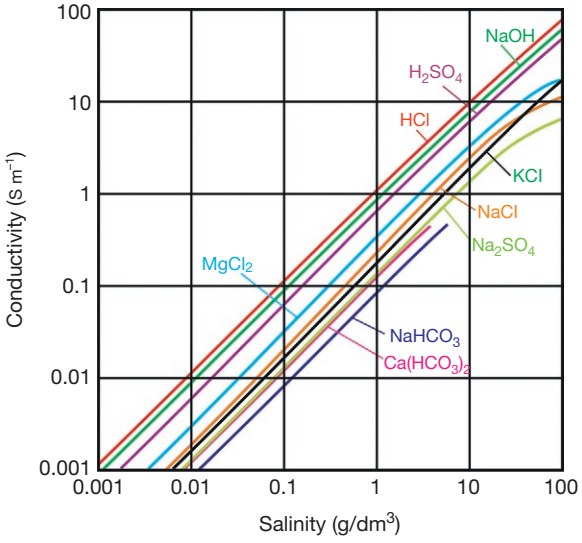


Figure 5 Conductivity of pore fluids as a function of salinity for different pore fluid compositions at 25 °C. Modified after Keller GV and Frischknecht FC (1966) *Electrical Methods in Geophysical Prospecting*, Pergamon Press, 517 pp.

In the upper crust, there are a number of ways in which conduction occurs. The most common is through the movement of charged ions in the pore fluid which either fully or partially saturates the pores of the rock. If the rock is composed of insulating minerals such as silicates, mineral conduction is negligible ($10^{-14} < \sigma_r < 10^{-10}$ S m^{-1}). By contrast, the conductivity of the pore fluids which they contain is much larger ($10^{-3} < \sigma_r < 1$ S m^{-1}). Hence, conduction takes place exclusively through the pore fluid by the movement of charged ions. Consequently, the conductivity of the rock is controlled by the way the pore fluids are connected throughout the rock, and this depends on the rock's microstructure.

The conductivity of the rock depends on

- the conductivity of the fluid that occupies the pores,
- the porosity of the rock,
- the degree to which the porosity is saturated with fluid, and
- the connectedness of pathways available for electrical conduction.

Each of these is discussed in turn:

Fluid conductivity. For most geomaterials, especially those at shallow levels in the crust where temperature and pressures are low, the matrix is effectively insulating and the bulk fluid provides the major contribution to the conductivity of the rock. Hence, we can say that the more saline the pore fluid, the higher the conductivity of any geomaterial saturated with it (Guéguen and Palciauskas, 1994).

Porosity. If the conductivity of a geomaterial depends upon conduction through the pores, it is intuitive that increasing the porosity will increase the electrical pathways available for conduction, and hence increase the electrical conductivity of the geomaterial (Guéguen and Palciauskas, 1994).

Saturation. A rock may have a large porosity, but the electrical conductivity of the whole rock may be very low if the pores are full of insulating fluid such as oil or gas. Hence, the degree to which the pores of the rock are saturated with a conducting fluid (usually brine) is also important, with higher conductivities occurring for higher water (brine) saturations (Guéguen and Palciauskas, 1994).

Connectedness. This is the last of the basic parameters affecting the conductivity of a geomaterial. Though critical to the final rock conductivity, it is unfortunately often overlooked. Rocks may have a large porosity that is completely saturated with a high-salinity fluid, yet have a low conductivity because the conducting pores are not connected (Guéguen and Palciauskas, 1994). This parameter is described by the connectivity of the geomaterial (Glover, 2009).

These four controlling parameters were brought together for the first time by Archie (1942), who developed a number of empirical laws to describe electrical conduction in reservoir rocks, which are described in subsequent sections of this chapter.

Another source of conduction involves conduction through the electrical double layer that is formed at the interface between the bulk fluid and the mineral surfaces (Adamson, 1976; Davis and Kent, 1990; Dukhin and Derjaguin, 1974; Glover et al., 1994; Hunter, 1981; Pride, 1994; Revil and Glover, 1997, 1998; Revil et al., 1999a,b; Sposito, 1989). This mechanism is more dominant at low bulk fluid salinities, where it provides more conduction than the weak bulk fluid, and also when the rock contains a significant fraction of clay minerals, which provide a greater surface area for surface conduction (see **Section 11.04.2.4**).

Surface conduction depends upon

- the specific conductance of the mineral surface/bulk fluid interface,
- a parameter describing the characteristic pore size of the rock, and
- the connectedness of surface conduction pathways.

Minor sources of conduction in the upper crust are associated with the presence of highly conducting phases such as graphite (Glover, 1996; Glover and Adám, 2008; Glover and Vine, 1992; Selway, 2014; Yang, 2011; Yoshino and Noritake, 2011) or sulfides (Duba et al., 1994; Einaudi et al., 2005; Glover and Vine, 1994), the latter of which rarely have sufficient connectedness to contribute significantly to the overall conductivity of the rock.

In particular areas, partial melting might be the cause of local high-conductivity anomalies (Glover et al., 2000b). Electrical conduction in melts is mediated by ions. The application of Archie's laws to these scenarios is difficult, but alternative methods now exist to allow there to be more than one conducting phase in a rock. These models are discussed in more detail in **Section 11.04.3.3**. Conduction through semiconducting minerals or through melt depends upon

- the conductivity of each of the phases present (melt, matrix, and conducting mineral),
- the volume fraction of each of the phases, and
- the electrical connectedness of each of the conducting phases.

In all of the cases listed earlier, we can note that the parameters that affect the electrical conductivity of the rock correspond to answers to the following questions:

- How conductive are the phases in the rock that conduct?
- How much of each phase is there?
- How connected is each of the phases?

11.04.2.3 Archie's Original Empirical Laws

Archie's laws were introduced in 1942 (Archie, 1942). We are introducing them here, before the section on mixing models, because of their importance, which was touched upon in the introduction.

It should be noted strongly that Archie's laws were originally empirical and had a small range of validity. However, it became increasingly clear that they could be applied successfully outside their initial range of validities and finally were proved analytically as an extension of the Lichtenecker and Rother (1931) mixing law (LR equation). The derivation can be found in a number of works including Glover (2010b). However, there are some questions concerning the theoretical derivation of the LR equation (Reynolds and Hough, 1957), and although the LR equation has recently been given a greater theoretical pedigree from effective medium approaches (Zakri et al., 1998), it would be fair to say that Archie's laws cannot be said to have an undisputed theoretical basis quite yet. However, recently, it has become clear that the simple Archie's laws are simply a special case of a more general mixing law that will be introduced in **Section 11.04.3.3**.

11.04.2.3.1 Formation factor and connectedness

Archie began by considering a porous medium of porosity ϕ, and naming the unitless ratio of the resistivity of the fully saturated rock ρ_o to that of the pore fluid saturating the pores ρ_f as the formation factor, F (which is sometimes called the resistivity formation factor):

$$F = \frac{\rho_o}{\rho_f} \quad [24]$$

(It should be noted that the symbols commonly used in the oil and gas industry are R_o in place of ρ_o, and R_w in place of ρ_f.)

The term formation factor was used because it was approximately constant for any given formation. The formation factor varies from unity, $F = 1$, which represents the case where $\rho_o = \rho_f$ (i.e., when $\phi \to 1$), and increases as the porosity decreases, with $F \to \infty$ as $\phi \to 0$. The formation factor can be less than unity, but only when the rock matrix is less resistive than the pore fluid and this is extremely rare (Glover, 2009).

The inverse of the formation factor is the ratio of the conductivity of the fully saturated rock to that of the fluid. This parameter is called the 'conductivity formation factor' or the 'connectedness' of the rock (G, no units) (Glover, 2009, 2010b). It has recently found utility in the extension of Archie's first law to n-conducting phases (**Section 11.04.3.3**). The symbol G should not be confused with conductance, which shares the symbol, but is rarely used in petrophysics, and is given the symbol G_c in eqn [6] earlier in this chapter to avoid confusion.

11.04.2.3.2 Archie's first law

The first experiments by Archie led him to the conclusion that the formation factor depends upon porosity in the form of an inverse power law (Archie's first law):

$$F = \phi^{-m} \quad [25]$$

with an exponent $-m$ (no units). He called the exponent the cementation exponent (factor or index) because he believed it to be related to the degree of cementation of the rock fabric. Higher values of m make the formation factor, and hence the rock conductivity, more sensitive to changes in porosity (Ellis and Singer, 2007). **Figure 6** shows the results of eqn [25].

The range of values for the cementation exponent is relatively small. A value of $m = 1$ is not observed for real rocks and represents a porous medium composed of a bundle of capillary tubes that cross the sample in a straight line. Rocks with a low porosity but a well-developed fracture network sometimes have cementation exponents that approach unity because the network has flow paths that are fairly direct. Here, we get the first taste that the cementation exponent has something to do with the connectedness of the pore and fracture network (where, for the time being, connectedness is considered to be a qualitative term for the general availability of pathways for transport). A cementation exponent equal to 1.5 represents the analytic solution for the case where the rock is composed of perfect spheres (Mendelson and Cohen, 1982; Sen et al., 1981). In fact, $m = 1$ (which is the trivial parallel conduction case) and $m = 1.5$ (Sen et al., 1981) were, until recently, the only two cases where an analytically derived value of the cementation exponent was known. A series of papers from 2004 onward have shown that Archie's law can be derived by applying continuum percolation theory to fractal porous media (e.g., Ewing and Hunt, 2006).

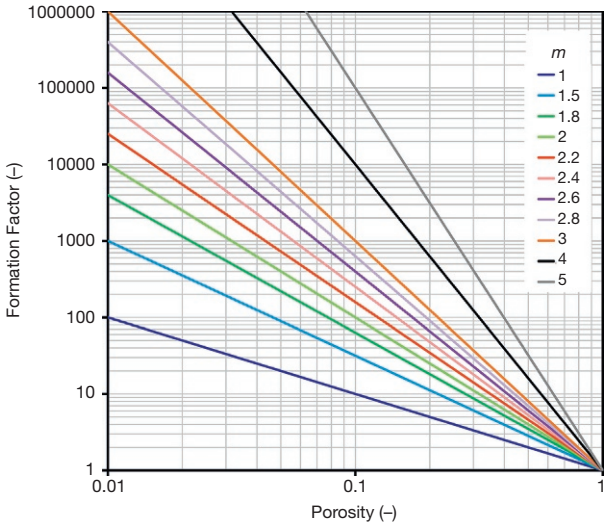


Figure 6 Formation factor as a function of porosity according to Archie's first law. The negative of the gradient of each curve provides the cementation exponent.

Most porous arenaceous sediments have cementation exponents between 1.5 and 2.5 (Glover et al., 1997). Values higher than 2.5, and as high as 5, are generally found in carbonates where the pore space is less well connected (Tiab and Donaldson, 2011). In general, the value of the cementation exponent increases as the degree of connectedness of the pore network diminishes, which rather supports it being called the cementation exponent. Incidentally, values of the cementation exponent less than unity are possible, arise when the matrix itself has a significant conductivity, and are observed in the modified Archie's law for two conducting phases (Glover et al., 2000a) (see **Section 11.04.3.3**).

The best method for obtaining a mean value of the cementation exponent of a set of rocks from the same formation is from the negative gradient of the graph of $\log(F)$ against $\log(\phi)$ (Tiab and Donaldson, 2011), which the oil and gas industry calls a Pickett plot, while values for individual samples can be found using

$$m = -\frac{\log(F)}{\log(\phi)} \quad [26]$$

Equations [24] and [25] are often combined under the name of Archie's law to give

$$\rho_o = \rho_f \phi^{-m} \quad [27]$$

which is an extremely useful form of the law. Here, $\rho_o \to \rho_f$ either as $\phi \to 1$ (i.e., the 'rock' is 100% pore fluid) or if $m = 0$ irrespective of the value of porosity. Even though $m = 0$ does not fall within the useful range for reservoir rocks, this result is a further indication that low values of cementation exponent represent good connectedness, and in the limit (i.e., when $m \to 0$), the connectedness of the conducting phase in the rock is optimal.

It should be noted that the correct use of Archie's first law requires accurate determinations of (i) the resistivity of the fully saturated rock, (ii) the resistivity of the fluid in the pores, and (iii) the relevant porosity for electrical transport

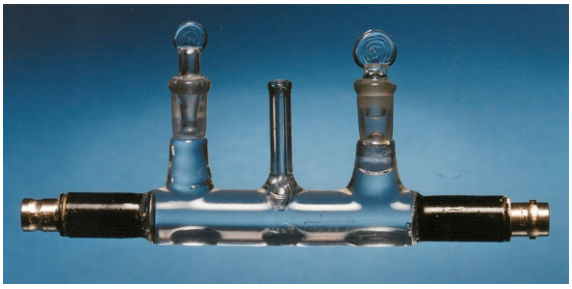


Figure 7 A cell for measuring fluid resistivity. Each end has a coaxial connection to a blacked platinum gauze electrode. There are two fluid ports for removing bubbles and a thermocouple well. The cell's geometric factor should be obtained by calibration with known fluids at each of the frequencies for which it will be used.

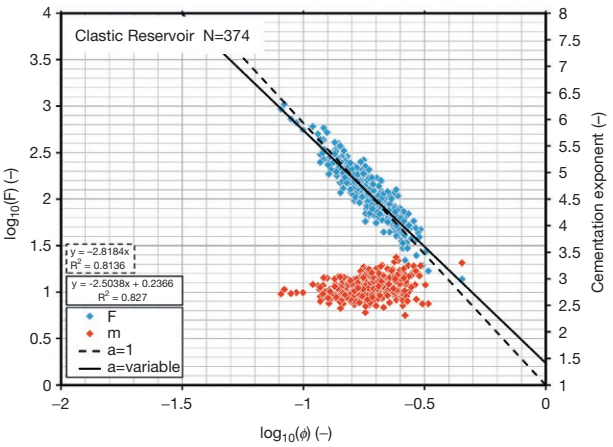


Figure 8 Formation factor (blue symbols) as a function of porosity for 374 samples from a single formation in a reservoir from the North Sea, UKCS. Two models have been fitted to the data. The dashed line is Archie's first law [25], while the solid line is Winsauer et al.'s (1952) modification with the '*a*' factor. Both seem to fit the data well according to the R^2 values, but the result is significantly different cementation exponents; Archie's first law [25] gives $m = 2.81814$, while Winsauer et al.'s modification gives $m = 2.5038$.

through the rock, which should be the effective porosity that contributes to the electrical conduction. Accurate determination of the resistivity of the fully saturated rock implies good control of the geometric factor that allows it to be calculated from the measured resistance, as well as certainty that the rock is fully saturated.

It is important that the resistivity of the brine in the pores is measured as there is a discrepancy between the salinity of brine in the stock saturating solution and that in equilibrium with the pores. At low salinities, the pore solution may be ten times more saline than that of the stock solution (Walker et al., 2014), leading to gross errors in the use of Archie's law if the values for the stock solution are used.

The resistivity of the solutions should be carried out in a specially designed cell with a known geometric factor (Figure 7) or a benchtop conductivity meter. The former is the better solution as it allows fluid resistivity measurements to be made at the same frequencies as the rock measurements if impedance spectroscopy is being performed, but the latter is acceptable providing the measurement is made in sufficient fluid with the electrode in the center of the beaker and a stirrer being employed. As with all electrical measurements, the temperature should be noted with the measurement, and corrections to a standard temperature carried out if required.

The accuracy of the porosity is more difficult. If one carries out porosity by helium pycnometry, by fluid saturation and by mercury injection porosimetry on the same sample, the measured porosities will be different in decreasing order. The question, that is still debatable, is which is the best measurement of porosity to use, especially as there are many other sources of porosity measurement too (i.e., by image analysis of thin sections, x-ray computed tomography (CT) and micro-CT scanning, positron emission tomography (PET) scanning, and nuclear magnetic resonance (NMR) spectroscopy)? At the moment, the most appropriate seems to be the fluid saturation approach. This approach at least has the advantage of employing fluids similar to those in the rock. It might be thought, therefore, that the same volume of pause might be saturated by using a reservoir fluid in the laboratory as occurred in the reservoir. At least, sufficiently similarly to make the measured porosity more relevant than that provided by other techniques.

It should be strongly noted that an extremely common form of Archie's first law contains an empirical factor a according to $F = a\phi^{-m}$ and hence $\rho_o = a\rho_f\phi^{-m}$. This empirical factor did not appear in Archie's original paper (Archie, 1942) and does not arise in the literature until Winsauer et al.'s (1952) and Wyllie and Gregory's (1953) papers. The use of the empirical factor is retained by many researchers because it improves the fit to sparse and low-quality datasets (Glover, 2012, pers. comm.). However, its use is analytically wrong because it implies that as the porosity tends to 100% (i.e., the sample has no matrix but is just the pore fluid), the sample has a resistivity $a\rho_f$. There is a contradiction here, because such a sample is solely pore fluid with a defined resistivity of ρ_f. Hence, the original equation is not generally valid unless the value of $a \equiv 1$. While the use of the empirical factor may have once been acceptable when Archie's first law had only an empirical pedigree, recent research indicates that nonunity values of a probably result from unresolved systematic errors in the porosity or resistivity measurements.

Figure 8 shows typical data from a reservoir in the North Sea. The data come from one well-defined sandstone formation. When the formation factor is plotted against porosity on logarithmic axes, two straight lines can be fitted. The dashed line represents Archie's first law [25], while the solid line incorporates the empirical factor a. In each case, the gradient represents the cementation factor m and it is instructive how different these are in the figure despite the goodness of fit R^2 value for the two fits being so similar. Here, the solid line provides the best assessment of the cementation exponent. In fact, for most data, the presence of the Winsauer a-factor allows a reasonably accurate cementation exponent to be calculated even if the data are not of the highest quality, and I would recommend its use in such cases, despite its non-analytical nature. Furthermore, the a-factor can be used as an indication of data quality (if close to unity, the quality of the data is good). The red points in Figure 8 show calculations of the cementation exponent on a plug-by-plug basis using eqn [26].

11.04.2.3.3 Resistivity index

Archie also examined the work of other investigators who did experiments on the resistivity of partially saturated sandstones. He observed that the bulk resistivity ρ_r of a rock that is partially saturated with an aqueous fluid of resistivity ρ_f is directly proportional to the resistivity of the rock when it is fully saturated with the same fluid, ρ_o:

$$I = \frac{\rho_r}{\rho_o} \qquad [28]$$

(Once again, it should be noted that the symbols commonly used in the oil and gas industry are R_o in place of ρ_o and R_t in place of ρ_r.)

The constant of proportionality I is called the resistivity index and describes the effect of partial desaturation of the rock. The resistivity index varies between unity and infinity depending upon the degree of saturation of the rock.

11.04.2.3.4 Archie's second law

Archie observed that the following relationship exists empirically for sandstones (Archie's second law):

$$I = S_w^{-n} \qquad [29]$$

where S_w is the fractional water saturation of the rock porosity and n is the saturation exponent.

The best method for obtaining a mean value of the saturation exponent of a sample in the laboratory is from the negative gradient of the graph of $\log(I)$ against $\log(S_w)$ (Tiab and Donaldson, 2011). This is a time-consuming measurement that involves the progressive desaturation of a core plug.

If we assume that a certain water saturation S_w is obtained by displacing some of the water from a previously fully saturated rock in a uniform manner, the volume fraction of water in the rock goes from ϕ to ϕS_w, and we can use eqn [27] to calculate the effective resistivity of the rock as $\rho_r = \rho_f(\phi S_w)^{-m}$ by assuming that the nonconducting fluid acts like the nonconducting rock matrix. Substituting this into eqn [28] and using eqn [27] again, but this time for the fully saturated rock, we get

$$I = \frac{\rho_r}{\rho_o} = \frac{\rho_f(\phi S_w)^{-m}}{\rho_f \phi^{-m}} = S_w^{-m} \qquad [30]$$

This is a demonstration of the analytic origin of Archie's second law, and predicts that the saturation exponent should be the same as the cementation exponent of the first law.

Equation [30], however, represents an oversimplification. The derivation of eqn [30] makes the assumption that we can replace ϕ in eqn [27] with ϕS_w while retaining the exponent $-m$. However, this is only possible if the connectedness of the phase ϕS_w is exactly the same as that of the previous phase ϕ. Of course, this is extremely unlikely to be the case. Imagine, for example, methane replacing water in a water-wet rock. The methane will preferentially occupy the center of the pores, and the water will preferentially occupy the pore space next to the mineral surfaces. The connectedness of the water in this partially saturated rock will be different, and generally less, than in the fully saturated case. Hence, we can say that $n \neq m$ generally because the connectedness of the fluid phase is a function of its volume fraction, and as the water is displaced from the pore space, the saturation exponent increases.

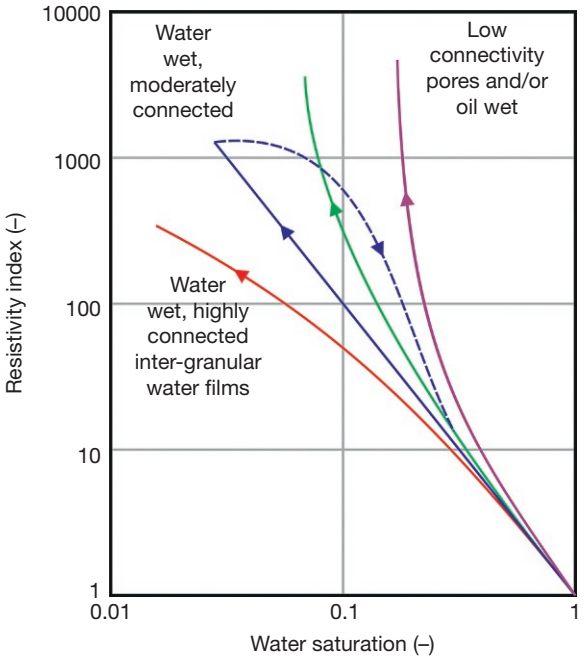


Figure 9 Resistivity index (as defined in eqn [28]) as a function of water saturation for a rock saturated with water and oil. Solid lines indicate drainage. The dark blue line, which represents Archie's second law ($n = 2$), is only valid for clean, water-wet moderately well-connected rocks. Rocks with higher connectivity due to the presence of intergranular water films tend to follow the red curve, while low connectedness or oil-wet rocks follow the green and purple trends. It should be noted that imbibition can often produce a nonlinear curve (dashed) due to the development of oil-wet patches in the rock.

The latest Archie's formulations that are generalized to take account of any number of conducting phases not only explicitly treat this problem (Glover, 2010b) but also show that both the first and the second Archie's laws are in fact the operation of the same general law but applied to a different frame of reference (see **Section 11.04.3.3**).

Since the laboratory determination of the saturation exponent is time-consuming and costly, it is common to hear that the saturation exponent has been taken to be equal to 2 in calculations. While it is true that there seems to be a strong preference for values of saturation exponent near 2 ± 0.5 for most water-wet rocks, oil-wet rocks show much higher values (4–5) (**Figure 9**) (Montaron, 2009; Sweeney and Jennings, 1960). Hence, it is extremely important to obtain accurate and representative values of this parameter for use in reserves calculations.

Once more, eqn [28] and eqn [29] can be combined to give

$$\rho_r = \rho_o S_w^{-n} \qquad [31]$$

Archie's second law can be expressed in graphical form, as shown in **Figure 9**. Equation [29] is shown by the straight blue line on this diagram. Clean rocks at high water saturations generally fall along such straight lines, with their slope representing the value of the saturation exponent. However, there also exist data that fall significantly off such lines. It is thought that data following the red line on the diagram represents rocks that have extremely high water connectedness due to the presence of intergranular water films, while data following the

green to purple trends (e.g., Sweeney and Jennings, 1960) represent low water connectedness particularly in oil-wet rocks. All of the solid curves are for drainage conditions. It has also been noted that it is possible to generate imbibition curves with different behaviors such as the dashed curve in the figure, which are believed to be caused by the development of patches of oil-wet rock that reduce the connectedness of the water phase and ensure that the saturation exponent is a function of the water saturation history of the rock (Montaron and Han, 2009).

11.04.2.3.5 Combined Archie's laws

The two Archie's laws can be combined into one equation:

$$\rho_r = \rho_f \phi^{-m} S_w^{-n} \quad [32]$$

This equation is of the greatest utility in the calculation of hydrocarbon reserves. In this application, values of porosity, pore fluid resistivity, and rock resistivity from borehole measurements are combined with the determinations of the cementation and saturation exponents, which are most often made in the laboratory on representative rock samples to calculate the water saturation. The water saturation can then be used with formation thickness, areal extent, and porosity to calculate the volume of hydrocarbons in the reservoir.

11.04.2.3.6 Electrical tortuosity and connectivity

In the discussion of Archie's laws, we have eluded to the importance of how well connected, or otherwise, a conducting phase, such as an aqueous fluid, is in the rock. This has traditionally been described by an electrical tortuosity. Electrical tortuosity τ_e is most often written as

$$\tau_e = F\phi \quad [33]$$

which, although useful does not show its origins clearly. It is possible to rewrite [33] as

$$F = \phi^{-1}\phi^{1-m} \quad [34]$$

where it can be recognized that the resistivity formation factor has a contribution arising purely from the presence of a scalar porosity and another that arises from the way that porosity is arranged, and which also depends upon porosity. This latter contribution is called the tortuosity, τ_e. Hence, the electrical tortuosity can also be expressed as

$$\tau_e = \phi^{1-m} \quad [35]$$

A full discussion of tortuosity can be found in the comprehensive review by Clennell (1997) and in the work of David (1993).

It should be noted that electrical tortuosity is not the same as hydraulic tortuosity, τ_h. Hydraulic tortuosity is defined as the mean flow path length between two points in a porous medium divided by the direct distance between those two points. The electrical tortuosity is often expressed as being equal to the square of the hydraulic tortuosity. Since there can be significant confusion between the two terms, it is always useful to define explicitly which type of tortuosity is meant, and to be certain which is being referred to when only the generic term 'tortuosity' has been used.

The inverse of electrical tortuosity is called the electrical connectivity χ. The conductivity formation factor G, which is sometimes called the connectedness, is related to the porosity and electrical connectivity by

$$G = \phi^m = \phi\chi \quad [36]$$

where it is clear that the overall connectedness of the rock depends on the porosity of a fully saturated pore space and its arrangement, the latter of which is given by the connectivity. Hence, the connectedness can be thought of as a measure of how the presence of the pores, when filled with a fluid of a given conductivity, controls the conductivity of the whole rock.

11.04.2.3.7 Validity of Archie's laws

Archie's laws came into being as empirical relationships that were formally valid only for the range of lithologies, temperatures, porosity, and pore fluids used to make the experimental measurements. However, very quickly, it became clear that the laws could be applied very successfully outside these initial ranges, and that has been confirmed by subsequent analytic derivations of Archie's first law.

Archie's laws, however, are subject to several restrictions:

- Archie's laws can only be used with a single conducting phase. Hence, the laws cannot be used at high temperatures in the presence of a melt because the rock matrix conducts significantly, or if there is a significant contribution from highly conducting accessory minerals. This restriction has been removed by the development of a version of Archie's first law for two conducting phases (Glover et al., 2000a,b,c) and then for any number of phases (Glover, 2010b).
- Archie's laws fail where the lithology contains minerals, usually shales, that provide a significant surface conductance and when the salinity of the pore fluids is low. Recent studies show that surface conduction effects can be very significant for a wide range of rocks and sediments especially at lower frequencies. However, Archie's laws can be used approximately in these cases if the pore fluid salinity is sufficiently high that it provides a conductance that is significantly larger than the surface conduction. Since the presence of shale surface conductance has always been important in the hydrocarbon industry, a number of modifications to Archie's first law have been proposed that are discussed in **Section 11.04.2.4** in the succeeding text, the most famous of which is by Waxman and Smits (1968).
- There is some question about the validity of Archie's laws at the extremes of the porosity range. At low porosities, there may be a fraction of the porosity ϕ_o that does not take part in conduction, which may be taken into consideration using a percolation limit approach (Bernabé and Bruderer, 1998; Hunt, 2004; Kirkpatrick, 1973; Seager and Pike, 1974; Shah and Yortsos, 1996):

$$F = (\phi - \phi_o)^{-m} \quad [37]$$

At high porosities, there is a dearth of data with which to check Archie's first law, but effective media models indicate that the rate of change of conductivity with porosity is less than that predicted by Archie's law.

11.04.2.4 Empirical Modifications for Shaley Rocks

The simple forms of Archie's laws assume that the only source of conduction in the rock is provided by the pore fluid and that ionic conduction in the pore fluid is uniform throughout the saturated pore space. While this is true for sedimentary rocks that are clay-free, it became clear that Archie's laws did not work for rocks with a significant fraction of clay, where there was an additional conduction contribution that was associated with conduction along clay surfaces. Initially, it was thought that atoms of Al^{3+} that substitute for Si^{4+} in the mineral structure would lead to a charge imbalance. Such an imbalance would then lead to an additional adsorption of cations onto the surface, which subsequently move and give rise to a surface conduction. We now know that a complex electrochemical structure exists at the interface between the mineral and the pore fluid, which is sometimes modeled as an electrical double layer (e.g., Revil and Glover, 1998) or an electrical triple layer (e.g., Revil and Leroy, 2004). Several types of surface conduction can occur within these structures, which are dependent on many factors, the most important of which are fluid salinity and pH (Parks, 1965; Revil and Glover, 1997, 1998; Revil et al., 1999a,b).

Clay minerals have a particularly high surface conduction. This was once thought to be, due to clays having a greater surface adsorption site density compared to, say, silica. However, it is now known that clay minerals generally have a lower surface adsorption sites density than many other minerals, but a much greater internal surface area that more than makes up for this. The effect of surface conduction is extremely important and is covered in **Section 11.04.4** in the succeeding text.

Archie's first law may now be modified to give

$$\sigma_r = \frac{\sigma_f}{F} + \frac{\sigma_s}{f} \qquad [38]$$

where we have used rock conductivity ($\sigma_r = 1/\rho_r$) and pore fluid conductivity ($\sigma_f = 1/\rho_f$) in place of the associated resistivities, together with the surface conductivity σ_s, and F and f are the formation factors associated with the bulk fluid conduction and the surface conduction, respectively. This equation assumes that the bulk conduction and the surface conduction occur independently and in parallel, giving rise to the simple addition, but does not assume that the flow paths are the same, hence the separate formation factors. However, it is usually assumed that both formation factors are the same and given the symbol F^* while rewriting [38] as

$$\sigma_r = \frac{1}{F^*}\left(\sigma_f + \frac{2\Sigma_s}{\Lambda}\right) = \frac{1}{F^*}(\sigma_f + BQ_v) \qquad [39]$$

The first equation in [39] arises from the formal definition of the formation factor (e.g., Revil and Glover, 1997; Schwartz et al., 1989), which will be discussed in **Section 11.04.3** in the succeeding text. The term Λ, which was introduced by Johnson et al. (1986), is a characteristic length (in m) associated with the pore structure of the rock and may be considered to be an effective pore radius. It is approximately equal to twice the pore volume divided by the pore surface area. The parameter Σ_s is called the specific surface conductivity or surface conductance (in S). It should be noted that the $2\Sigma_s/\Lambda$ term depends upon an assumption that is explicit in the definition of Λ, which is that the interfacial conductivity extends only a short distance into the pores, which is the same as saying that the pore walls appear flat with respect to the thickness of the electrical double layer (i.e., a thin double-layer assumption) (Johnson et al. 1986).

Applying an analogy of [10] in two dimensions, we obtain $\Sigma_{si} = n_{si}Z_i eB_i$ for each of i types of ion on the surface, where Σ_{si} is the specific surface conductance of each type of surface ion, n_{si} is the surface number density of each type of surface ion, Z_i is the valency of each type of surface ion, and B_i is the surface mobility of each type of surface ion. Now, taking the mean behavior of all ions and including $\Lambda = 2V_p/S_p$ allows us to write $2\Sigma_s/\Lambda = BQ_v$, where $Q_v = n_s Z e S_p/V_p$ is the excess surface charge taking part in surface conduction per unit pore volume (Guéguen and Palciauskas, 1994).

It should be noted that the excess charge per volume is related to the traditional measurement of CEC. The CEC measurement (e.g., Ross and Ketterings, 2011) provides a chemical measurement of excess charge per weight of the sample. The value is commonly given in milliequivalents per gram (mequiv. g^{-1}) where, for NaCl solutions, 1 mequiv. $g^{-1} = 96\,320$ C kg^{-1} (e.g., Patchett, 1975). While CEC measurements are carried out on crushed samples, and one could question their validity for whole rock samples, it is possible to take the CEC measurement for a rock and calculate the equivalent value of Q_v providing the density and porosity of the rock are known.

The right-hand side of [39] was developed by Waxman and Smits (1968) and used as the standard shaley sand model for conductivity for many years. **Section 11.04.3.5** discusses models for shaley sands in more detail and compares the Waxman and Smits model with others that are more suited for application in the hydrocarbon industry.

11.04.2.5 The Range of Electrical Conductivity of Near-Surface Rocks

Huge numbers of experiments have been carried out on rocks from the upper crust in the last 70 years; many of them by the oil and gas industry. While most remain confidential, it is possible to get an extremely good idea of the range of electrical conductivities from extremely low-conductivity microgranites with vanishingly small porosities and almost no microfracturing (e.g., Ailsa Craig, which is still used to make curling stones), through reservoir rocks where the conductivity is controlled by the fluids in the pore spaces and clays, to veins of highly conducting minerals such as sphalerite or graphite. **Table 1** shows some typical ranges of conductivity and resistivity obtained from the literature, while some of the same information is available in **Figure 1**. It is not within the remit of this chapter to review the huge number of measurements that are available. However, good compilations exist in the literature, the foremost of which are by Keller (1989), while more data can be found in Schön (2004) and to a lesser extent in Guéguen and Palciauskas (1994).

A great number of processes have the potential to affect the conductivity of rocks. **Box 1** shows a compilation of the main processes and the direction in which each one influences the conductivity of the geomaterial.

Table 1 A summary table of the conductivities and resistivities of common rocks and minerals (at surface temperatures and pressures)

Rock	*Conductivity (S m^{-1})*	*Resistivity (Ωm)*
Granite	0.001–0.00001	1000–100 000
Gabbro	0.001–0.00001	1000–100 000
Basalt	0.1–0.0001	10–100 000
Andesite	0.001–0.00001	1000–100 000
Sand and gravel	0.05–0.002	20–5000
Sandstone	0.1–0.001	50–1000
Limestone	0.002–0.00001	5000–100 000
Dolomite	0.001–0.0002	1000–50 000
Conglomerate	0.01–0.0001	100–10 000
Coal	0.1–0.001	10–1000
Shales	0.001–1	50–1000
Clays	0.05–0.01	20–100
Graphite	100–0.01	0.01–100
Massive sulfides	1000–1	0.001–1
Salt water	10–1	0.1–1
Brackish water	1–0.1	1–10
Freshwater	0.1–0.01	10–100
Permafrost	0.02–0.00001	500–100 000
Sea ice	0.05–0.001	20–1000

Modified after Palacky GV (1987) Resistivity characteristics of geologic targets. In: *Electromagnetic Methods in Applied Geophysics, 1. Theory*, p. 1351

Box 1 Effect of various processes upon the conductivity of geomaterials

Effect	Conductivity
Increasing porosity	↑
Increasing pore fluid salinity	↑
Increasing surface conductance	↑
Increasing connectedness	↑
Increasing connectivity	↑
Increasing cementation exponent	↓
Increasing clay alteration	↑
Dissolution	↑
Shearing/faulting	↑↓
Weathering	↑
Induration	↓
Carbonate precipitation	↓
Silicification	↓
Metamorphism	↑↓
Compaction	↓
Overpressure	↑
Increasing formation factor	↓
Increasing tortuosity	↓
Hydrocarbon production/imbibition	↑
Hydrocarbon emplacement/drainage	↓
Increasing saturation exponent	↓
Increasing thermal/microfracturing	↑
Increasing oil wettability	↓
CCS emplacement	↓
Hydrofracturing	↑
Formation temperature	↑
Increasing overburden pressure	↓

11.04.3 Mixing Models for the Electrical Conductivity of Geomaterials

Mixing models are incredibly important for the measurement and calculation of the electrical properties of all complex and porous media. In this section, we will review some of the main mixing models for the steady-state electrical conductivity of geomaterials, pointing out their advantages and disadvantages. Each can be written in terms of resistivities. Some of the models in this section are theoretical, while many are empirical and some empirical models have an increasingly good theoretical pedigree. Mixing models for the permittivity and complex electrical properties of rocks are left until Section 11.04.5.5.

The models fall into three main groups, which are described in each of the following subsections. A special subset of mixing models that includes variable water saturations has also been considered. Many of the models referred to in this section are presented in Table 2 and are graphed in Figure 10.

11.04.3.1 Geometric Models

These models consider two, three, or more conducting phases and build the arrangement of those phases into the model itself. In other words, these are geometric models that are only valid if the arrangement of the conducting phases approximates to that of the model (first part of Table 2).

The parallel model can have many conducting phases with different phase conductivities σ_i and different phase volume fractions ϕ_i, but they all have to be arranged electrically in parallel. The equation for the resulting conductivity is the same as an arithmetic mean weighted by the volume fractions. By contrast, the perpendicular model considers the conductivities to be arranged in series and is expressed by a weighted harmonic mean. Here, the current flows perpendicularly to the layers of conducting phase crossing each of them in series. Clearly, these are extreme cases and the actual conductivity of a rock would fall somewhere in between these bounds.

In 1962, Hashin and Shtrikman (1962) produced two other bounding models that are more restrictive and hence more useful than the parallel and perpendicular cases. These are called the Hashin and Shtrikman upper and lower bounds or HS+ and HS−, respectively. Unlike the series and parallel models, however, the Hashin and Shtrikman bounds are only valid for two conducting phases. There is currently active, and extremely mathematical, research into the extension of the HS bounds to multiple phases and multiple dimensions (e.g., Liu, 2010).

Ten years after the publication of the Hashin and Shtrikman bounds, Waff (1974) produced a model that shares many of the basic structural qualities of a granular rock in that its two phases were considered to be in the form of conducting spheres with conducting spherical shells around them, which approximates to the scenario where grains are surrounded by a pore fluid. It turned out that the equation he produced is mathematically the same as the Hashin and Shtrikman upper bound, giving it more validity for use in granular porous media with subrounded grains.

The first part of Figure 10 shows a comparison between the parallel, perpendicular, HS+/Waff, and HS− models for a two-phase mixture with the first phase having the nominal conductivity of 0.01 S m^{-1} and the second having the nominal

Table 2 Some of the most common mixing models for electrical conductivity in porous media

Name	Conducting phases: Min	Conducting phases: Max	Equation	References	Notes
Models without variable exponents					
Parallel model	1	Many	$\sigma_{\text{eff}} = \sum_{i=1}^{n} \phi_i \sigma_i$	Guéguen and Palciauskas (1994), Luo et al (1994), Somerton (1992)	Arithmetic mean. Parallel layers of constant arbitrary thickness with conductivity σ_i arranged axially to current flow
Perpendicular model	1	Many	$\frac{1}{\sigma_{\text{eff}}} = \sum_{i=1}^{n} \frac{\phi_i}{\sigma_i}$	Guéguen and Palciauskas (1994), Luo et al (1994), Somerton (1992)	Harmonic mean. Parallel layers of constant arbitrary thickness with conductivity σ_i arranged normally to current flow
Random model	1	Many	$\sigma_{\text{eff}} = \prod_{i=1}^{n} \sigma_i^{\phi_i}$	Guéguen and Palciauskas (1994), Luo et al (1994), Somerton (1992)	Geometric mean. Arbitrary shaped and oriented volumes of conductivity σ_i distributed randomly
Hashin–Shtrikman upper bound	2	2	$\sigma_{\text{eff}}^{+} = \sigma_2 \left(1 - \frac{3(1-\phi_2)(\sigma_2-\sigma_1)}{3\sigma_2 - \phi_2(\sigma_2-\sigma_1)}\right)$	Hashin and Shtrikman (1962)	Commonly denoted HS+. Derived from effective medium considerations
Hashin–Shtrikman lower bound	2	2	$\sigma_{\text{eff}}^{-} = \sigma_1 \left(1 + \frac{3\phi_2(\sigma_2-\sigma_1)}{3\sigma_1 + (1-\phi_2)(\sigma_2-\sigma_1)}\right)$	Hashin and Shtrikman (1962)	Commonly denoted HS−. Derived from effective medium considerations
Waff model	2	2	$\sigma_{\text{eff}} = \frac{\sigma_2 + (\sigma_1-\sigma_2)(1-(2\phi_2/3))}{1 + (\phi_2/3)(\sigma_1/\sigma_2 - 1)}$	Waff (1974)	Based on concentric spheres of varying sizes with volume of core (fractional volume of phase 1) to volume of shell (fractional volume of phase 2) ratio constant. Functionally equivalent to HS+
Brick-layer model	2	2	–	Beekmans and Heyne (1976)	Cubes of one phase arranged in a three-dimensional grid with a second phase occupying the gaps between the blocks. Only valid for small percentages of interblock phase
Modified brick-layer model	2	2	$\sigma_{\text{eff}} = \frac{\sigma_2\left(\sigma_2\left(\phi_1^{2/3}-1\right) - \sigma_1\phi_1^{2/3}\right)}{\sigma_1\left(\phi_1 - \phi_1^{2/3}\right) + \sigma_2\left(\phi_1^{2/3} - \phi_1 - 1\right)}$	Schilling et al. (1997)	Modified to allow validity to be extended to cover the range $0.00 \le \sigma_1 \le 1.00$ (0–100%) Almost coincident with HS+
Models with variable exponents					
Lichtenecker–Rother equation	1	2	$\sigma_{\text{eff}} = (\sigma_1^{1/m}(1-\phi_2) + \sigma_2^{1/m}\phi_2)^m$	Lichtenecker and Rother (1931) Korvin (1982)	Derived from the theory of functional equations under appropriate boundary conditions. Formally, the same as Archie's law if $\sigma_1 = 0$
Bussian's equation	2	2	$\sigma_{\text{eff}} = \sigma_2\phi_2^m \left(\frac{1-\sigma_1/\sigma_2}{1-\sigma_1/\sigma_{\text{eff}}}\right)^m$	Bussian (1983)	Derived from effective medium theory
Conventional Archie's law	1	1	$\sigma_{\text{eff}} = \sigma_2\phi_2^m$	Archie (1942)	Derived empirically, but provable analytically for special cases
Extended and generalized models with variable exponents					
Lichtenecker–Rother equation (generalized)	1	Many	$\sigma_{\text{eff}} = \left(\sum_{i=1}^{n} \sigma_i^{1/m}\phi_i\right)^m$	Lichtenecker and Rother (1931) Korvin (1982)	Logical extension of the Lichtenecker and Rother model made in this work
Modified Archie's law	1	2	$\sigma_{\text{eff}} = \sigma_1(1-\phi_2)^p + \sigma_2\phi_2^m$ where $p = \frac{\log(1-\phi_2^m)}{\log(1-\phi_2)}$	Glover et al. (2000a)	Derived from conventional Archie's law by considering boundary conditions implied by geometric constraints
Generalized Archie's law	1	Many	$\sigma = \sum_{i}^{n} \sigma_i \phi_i^{m_i}$ where *exact solution* $m_j = \log\left(1 - \sum_{i \neq j} \phi_i^{m_i}\right) \Big/ \log\left(1 - \sum_{i \neq j} \phi_i\right)$ *first order approximation* $m_j = \sum_{i \neq j} \phi_i^{m_i} \Big/ \sum_{i \neq j} \phi_i$	Glover (2010b)	Derived from the conventional Archie's law by considering boundary conditions implied by geometric constraints

Notes. There is either 1 phase, 2 phases, or n phases each denoted by subscripts, with the subscript i denoting the ith phase in an n-phase set. Conductivities are noted as σ, while cementation and analogous exponents are denoted by m or p. The volume fraction of each phase is denoted by ϕ.

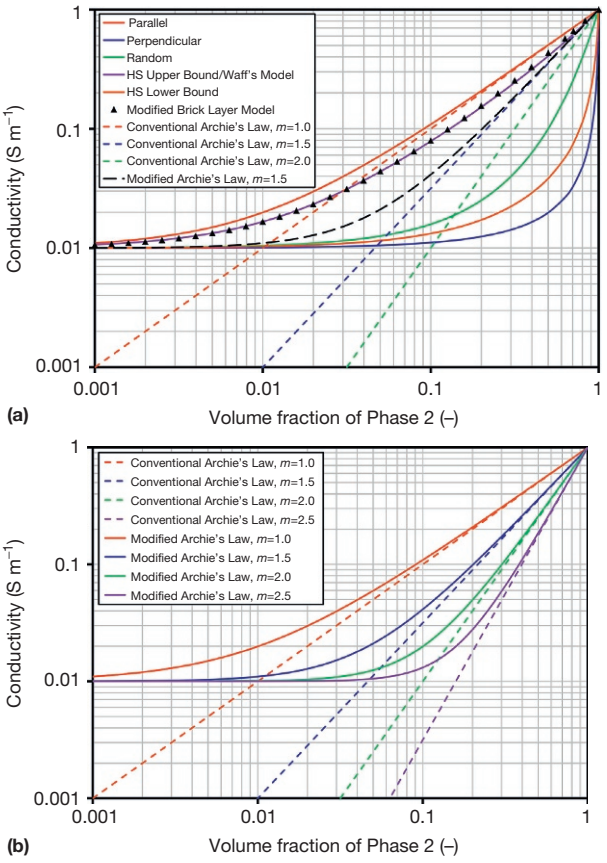


Figure 10 (a) Comparison of various mixing models as a function of the volume fraction of Phase 2. In this generic diagram conductivity of Phase 1 is $\sigma 1 = 0.01$ S m^{-1} and that for Phase 2 is $\sigma 2 = 1.0$ S m^{-1}. Geometrical models are shown by solid lines, except the modified bricklayer model which is shown by triangles because it is so similar to the HS upper bound, which it overlies. The three Archie's law implementations, for $m = 1.0$, 1.5, and 2.0, respectively, are shown by short dashed lines. The single implementation of the modified Arches law for m equals 1.5 is shown in black with a long dashed line. Note that the parallel and perpendicular models describe the extreme bounds of the conductivity for any given volume fraction of Phase 2, while the HS upper and lower bounds are more restrictive. (b) Comparison of the performance of the conventional Archie's law (Archie, 1942) with the modified Archie's law for two conducting phases (Glover et al., 2000). The conventional Archie's law is given by short dashed lines, while the modified Arches law by solid lines. Note that in no cases does the conventional Archie's law model the conductivity well at low volume fractions of Phase 2, while the modified Archie's law converges on the conductivity of Phase 1 as the volume fraction of Phase 2 goes to 0. In each case the modified Archie's law becomes asymptotic with the conventional Archie's law at large volume fractions of Phase 2. Reproduced from Glover PWJ, Hole MJ, and Pous J (2000). A modified Archie's law for two conducting phases. Earth and Planetary Science Letters 180: 369-383. http://dx.doi.org/10.1016/S0012-821X(00)00168-0.

conductivity 1 S m^{-1}, as a function of the volume fraction of the second phase. It is clear that the HS+ and HS− bounds fall between the values for the parallel and perpendicular models.

Another approach is to consider cubic bricks instead of spheres in a so-called brick-layer model (BLM). Beekmans and Heyne (1976) did this to produce a model for two phases with different conductivities and volume fractions. They considered the blocks to be solid grains with a given conductivity and the material between the blocks to be melt of a higher conductivity. Unfortunately, this model was only valid for the volume fraction of the interblock phase that amounted to less than a few percent (Barsoukov and Macdonald, 2005). The model was modified by Schilling et al. (1997) to be valid for all interblock volume fractions and is known as the modified brick-layer model (MBLM). Schilling et al. then used the MBLM to model the electrical conductivity of partial melting in laboratory tests very successfully (Partzsch et al., 2000), which implied that the structure of the melt and solids in their experimental samples must have been approximated by the blocky structure of the model. Plotting the MBLM in Figure 10 shows it to be extremely similar to HS+, but not exactly the same.

Finally, an interesting model is based upon the weighted geometric mean. Although it cannot be proven analytically, this equation has been shown to arise naturally from a random three-dimensional mixture of phases of different conductivities σ_i and different phase volume fractions ϕ_i. One would think that such an equation would not be useful in the modeling of electrical conductivity in geomaterials, but it has already found application in calculating the permeability of a rock with different fractions of different grain sizes using electrical data (Glover et al., 2006a,b). Figure 10 shows that the geometric mean (labeled 'random model' in Figure 10) is closer to the Hashin and Shtrikman lower bound (HS−) than the MBLM or Waff model.

The advantage of the geometric models is that in many cases, one can treat many different conducting phases. Their disadvantage is that they represent only one particular geometry.

11.04.3.2 Models with Variable Exponents

The second part of Table 2 contains models with variable exponents.

The advantage that Archie's law and the other models in this subsection have over the geometric models is that they contain one or more variable exponents. These exponents are a measure of the connectedness of each of the conducting phases in the rock. The exponents not only give an extra degree of flexibility when fitting the equation to data but also can be inverted to provide the exponents, which can then be used as a measure of the connection of each phase in the rock.

The disadvantage with the original Archie's law is that only one conducting phase is allowed. This becomes clear in the top part of Figure 10, where each of the curves for Archie's cementation exponents $m = 1$, 1.5, and 2, respectively, tends to zero as the volume fraction of the second (conductive) phase tends to zero. If one were modeling a system with a conducting matrix (Phase 1) and a conducting fluid (Phase 2), the model would fail, particularly at low porosities and for low-conductivity fluids. Figure 10(a) shows such a scenario, where the conductivity of the matrix (Phase 1) is 0.01 S m^{-1} and that of the conducting fluid (Phase 2) is 1 S m^{-1}. Use of the conventional Archie's law as represented by the straight lines in the figure leads to errors. If we ask the conventional Archie's model to tell us what the conductivity of a rock composed of 1% fluid and 99% matrix is, we get 0.01, 0.001, and 0.0001 S m^{-1} for $m = 1$, 1.5, and 2, respectively. In all cases, the calculated conductivity is equal to or significantly lower than the conductivity of the rock matrix even though the rock matrix makes up 99% of the rock. This is an error that the geometric models do not make, tending to the conductivity of the matrix as the other phase

decreases to zero (Figure 10(a)). It is also an error that the modified Archie's law is not subject to (Section 11.04.3.3).

Although Archie's first law is probably the simplest of this group and was, as we have already seen, originally derived empirically, the Lichtenecker and Rother (1931) equation (LR equation) predated it by ten years. The LR equation is also more general, being valid for two or more phases, and indeed collapses to become formally the same as Archie's first law if all but one of the phase conductivities become zero. Initially, the LR equation was also empirical but has recently gained more theoretical pedigree (Zakri et al., 1998). It is also interesting in that in the LR equation, the conductivity terms have the exponents, not the phase fractions. Bussian's (1983) equation was derived from effective medium theory and was thought to be a good model for the complete range of parameters, but it has been shown recently that it diverges from experimental results at low fluid conductivities. In addition, it is difficult to solve because it is nonlinear. However, a conformal mapping code is now available to simplify the process (Glover et al., 2010).

11.04.3.3 Extended and Generalized Archie's Laws

In 2000, Glover et al. (2000b) needed to model the electrical conductivity of a rock where the solid grains had a finite conductivity because they were at high temperature but there was also a high-conductivity melt fraction of unknown geometry. Under these conditions, the modified brick-layer model cannot be used. Instead, they produced an extended version of Archie's law for two conducting phases where each phase had an exponent that described its degree of connection, and the two exponents were related to each other Glover et al. (2000a). This, so-called, modified Archie's law (MAL) not only has two conducting phases, but also allows the electrical connectedness of the two phases to be varied, so different geometries can be modeled.

The derivation of modified Archie's law is heuristic and makes an important assumption that electrical conduction in each of two conducting phases that completely describe the porous medium can be considered independently. In other words, any mathematical description does not contain a term describing the interaction between the phases. Formally, such a term is written as a Stieltjes integral. That the conduction in the two phases can be considered independently does not of course imply that the two conduction paths are independent. Conduction may start off in Phase 1 and cross over into Phase 2 and back again many times, just that when the conduction occurs in Phase 1, it is governed by Phase 1 properties and likewise for Phase 2. When current passes from Phase 1 to Phase 2, whatever perturbation that has on the conductivity of the whole rock is canceled out by the currents passing from Phase 2 to Phase 1 in other parts of the rock. It is hypothesized that this canceling of the interaction terms is fundamentally a consequence of Gauss' law. The result is that each of the phases can be described by an Archie-like term:

$$\sigma_r = \sigma_1 \phi_1^p + \sigma_2 \phi_2^m \qquad [40]$$

where there are two phases (subscripted 1 and 2), each phase has a volume fraction ϕ_1 and ϕ_2, respectively, and each has an intrinsic conductivity σ_1 and σ_2, respectively. Phase 1 has an exponent p and Phase 2 has an exponent m, using the symbology of the original paper (Glover et al., 2000a). In the usual interpretation, Phase 1 is the matrix and Phase 2 is the fluid that occupies the pores of the rocks; hence, eqn [40] becomes

$$\sigma_r = \sigma_{\text{matrix}}(1-\phi)^p + \sigma_f \phi^m \qquad [41]$$

where ϕ is the porosity of the porous medium, σ_{matrix} is the conductivity of the rock matrix, and σ_f is that of the fluid occupying the pores. The application for which the modified Archie's law was developed, however, did not have an aqueous pore fluid. For that application, Phase 1 was hot crustal rock whose conductivity was described using equations of the form of eqn [11] for continental crustal lithologies, and Phase 2 was melt (Glover et al., 2000b).

The modified Archie's law has two exponents, each of which describes the electrical connectedness of its phase, and this allows a range of behaviors including that which approximates to the HS upper bound/Waff model ($m = 1.15$–1.25) and the HS lower bound ($m = 1.8$–3). Table 3 shows a full set of approximations of the modified Archie's law to various other models. The first part of Figure 10 includes one curve using the MAL with a cementation exponent of the second phase equal to 1.5. It is interesting to note that the MAL behaves in general like the geometric models, but its position between the parallel and the perpendicular models is defined by the cementation exponent of the second phase. This behavior is shown clearly in the second part of Figure 10, which compares the behavior of the conventional and modified Archie's laws for four values of cementation exponent.

Table 3 Approximation of the modified Archie's law (MAL) for various exponents to other models

Approximate modified Archie's law exponents	
Model	m
Parallel model	1
Perpendicular model	3–5
Random model	1.7–2.2
Hashin–Shtrikman upper bound/Waff model	1.15–1.25
Hashin–Shtrikman lower bound	1.8–3
Modified brick-layer model	1.15

Where there is a range given, the lower values produce a good fit at low values of the volume fraction of the phase with the higher conductivity, and the higher values in the table produce better fits for high values of the volume fraction of the phase with the higher conductivity.

The two exponents in the modified Archie's law are not independent of each other. Just in the way that the volume fractions of the two phases always sum to unity, there is a finite connectedness available to the three-dimensional rock. This means that if one phase, say, Phase 1, increases its connectedness (i.e., the exponent p decreases), the connectedness of Phase 2 must increase (i.e., the exponent m increases), and vice versa. The two exponents are linked by the relationship

$$p = \frac{\log\left(1-\chi_2^m\right)}{\log\left(1-\chi_2\right)} \qquad [42]$$

The modified Archie's law has subsequently been used successfully from concrete to food engineering.

The question arose naturally whether a form of Archie's law for n-conducting phases existed. Following the development of the concept of connectedness in 2009 (Glover, 2009), such a law was formulated and partially validated using a numerical

approach (Glover, 2010b). This general law will probably only rarely find an application. However, interestingly, the generalized Archie's first law automatically generates the second Archie's (saturation) law and seems not to require the concept of percolation in order to be a complete description of the electrical conductivity of an *n*-phase material, geologic or not.

11.04.3.4 Empirical Models for Full Water Saturation

In the literature, there exist a large number of models that have been developed empirically for a particular company and even a particular oil or gas field. Once successful in a given field for a given company, they are often applied elsewhere without consideration of their validity. These equations should generally be avoided.

Company models are generally the same form as Archie's law, but with particular values for a and m. Examples include the 'Humble' formula ($a=0.62$, $m=2.15$), which was developed for the Humble Oil Company by Winsauer et al. (1952) from 30 samples (28 sandstones, 1 limestone, and 1 unconsolidated sandstone) and the Tixier equation ($a=0.81$, $m=2$), which uses the same data and forces $m=2$. Other equations which are based on more data include that of Carothers and his colleagues who used a database of 793 sandstone samples ($a=1.45$, $m=1.54$) and the Shell formula, which was developed for low-porosity (<9%) unfractured carbonates ($a=1$, $m=1.87+0.019/\phi$). This model is one of the many that use an empirical relationship to obtain the cementation exponent and then use that cementation exponent within Archie's law. Others include Focke and Munn (1987), Nugent et al. (1978), Nurmi and Frisinger (1983), and Rasmus (1983).

These models were often not intended for general application, but have subsequently been applied generally. They have no scientific merit in themselves, and most of them represent specific cases of Archie's first law. It is best to regard them as obsolete and use one of the other more appropriate models given in Table 2.

11.04.3.5 Shaley Sand Water Saturation Equations

Most shales have moderately low resistivities, which can be seen on any resistivity log, and log analysts explain the fact by attributing the extra conduction to absorbed waters on clay surfaces. This is a simplistic way of explaining effects that we now describe using electrical double- and triple-layer models (e.g., Revil and Glover, 1997, 1998; Revil and Leroy, 2004).

The presence of clays as a component of shaley sandstones introduces an extra conductivity contribution that should be combined with the conductivity of the pore formation water in detailed reservoir analysis. If no account is taken of the extra conductivity in a shaley sand zone by using the simple combined Archie's equation [32] to calculate water saturation, the result will be an overestimation of water saturation, since the resistivities have been reduced below their true values by the conductivity of the shale component. Hence, potentially economic hydrocarbon-bearing intervals may be missed.

There are two main families of resistivity model for partially saturated shaley sandstones. Both have developed from Archie's equations and accommodate the conductivity effects of clay minerals in order that more accurate water saturations may be computed (Worthington, 1985). All the models consider that the shale is a homogeneous conductive medium and add the conductivity of the shale fraction (V_{sh}) to Archie's law component of conductivity. These equations have the form

$$\sigma_t = \sigma_{clean} + \sigma_{shale} \quad [43]$$

where the total conductivity of the rock σ_t is composed of that which it would have if it were clay-free σ_{clean} plus that associated with the shales σ_{shale}. Although the physical basis of these models is incorrect because the shale and pore fluid contributions do not, in general, occur in parallel and independently, the equations often provide approximate solutions to water saturation, especially when the equation parameters are adjusted so that the results conform to local water saturation data measured from cores or production tests. Table 4 shows some of the main models of this type.

All but two of the models in Table 4 assume that pore fluid-mediated conduction and shale-mediated conduction (whatever its physical mechanism) are independent, occur in parallel, and hence can be described by a sum of conductivities. One would expect that these models perform better in laminar or layered shales. The two remaining models (Waxman and Smits (1968) and the dual-water model (Clavier et al., 1977; Coates et al., 1983)) are also formed by a sum of conductivities, which would imply parallel conducting mechanisms, but their use of different formation factors allows them to also perform well where the clay is distributed throughout the rock in a dispersed manner or is part of the rock structure.

It has been pointed out that none of these models are correct in a theoretical sense (Doveton, 2001), but in approximating the theoretical behavior, some can provide fairly accurate solutions for the hydrocarbon industry. Some of these equations are still widely used because of their relative simplicity and limited demands on additional input parameters.

One of the most commonly used equations of this type is the Simandoux equation (Simandoux, 1963). The original equation contained a factor ε, where $\varepsilon=1$ when $S_w=1$ and $\varepsilon<1$ when $S_w<1$. In fact, it is usually the modified Simandoux equation (Bardon and Pied, 1969) that is used these days. In this model, $\varepsilon=S_w$. The modified Simandoux equation can be written in a generalized polynomial form that makes it easier to implement numerically and accounts for saturation exponents other than $n=2$:

$$S_w^n + \left[\frac{V_{sh}\rho_w\phi^{-m}}{\rho_{sh}}\right]S_w - \frac{\rho_w\phi^{-m}}{\rho_t} = 0 \quad [44]$$

It is known that the Simandoux equation tends to underestimate the water saturation in oil reservoirs (Doveton, 2001; Smith and Rouleau, 1977). The underestimation of water saturation leads to an overestimation of the oil saturation that Doveton (2001) has suggested has been a possible reason for the popularity of the equation.

The Simandoux and Bardon and Pierre equations make a number of assumptions. First, the clean and shaley contributions can be considered to be independent and arranged electrically in parallel, leading to the conductivities being summed. Second, the shale fraction has a very good connectedness for both the shale volume and the water associated with that shale, as indicated by the unity exponents for V_{sh} and S_w in the second term (cf. Table 4). Examination of the other equations of this type shows the first assumption also applies to them, while the second assumption applies to most. The Hossin

Table 4 A selection of shaley rock water saturation models (arranged in chronological order)

Models	*Equation in terms of resistivity*	*Equation in terms of conductivity*	*Reference*
Poupon et al.	$\frac{1}{\rho_t}=\frac{(1-V_{sh})S_w^2}{F\rho_w}+\frac{V_{sh}}{\rho_{sh}}$	$\sigma_t=\frac{(1-V_{sh})S_w^2}{F}\sigma_w+V_{sh}\sigma_{sh}$	Poupon et al. (1954)
Hossin	$\frac{1}{\rho_t}=\frac{S_w^2}{F\rho_w}+\frac{V_{sh}^2}{\rho_{sh}}$	$\sigma_t=\frac{S_w^2}{F}\sigma_w+V_{sh}^2\sigma_{sh}$	Hossin (1960)
Simandoux	$\frac{1}{\rho_t}=\frac{S_w^2}{F\rho_w}+\frac{\varepsilon V_{sh}}{\rho_{sh}}$	$\sigma_t=\frac{S_w^2}{F}\sigma_w+\varepsilon V_{sh}\sigma_{sh}$	Simandoux (1963)
Waxman and Smits	$\frac{1}{\rho_t}=\frac{S_w^2}{F^*\rho_w}+\frac{BQ_vS_w}{F^*}$	$\sigma_t=\frac{S_w^2}{F^*}\sigma_w+\frac{BQ_vS_w}{F^*}$	Waxman and Smits (1968)
Modified Simandoux	$\frac{1}{\rho_t}=\frac{S_w^2}{F\rho_w}+\frac{V_{sh}S_w}{\rho_{sh}}$	$\sigma_t=\frac{S_w^2}{F}\sigma_w+V_{sh}\sigma_{sh}S_w$	Bardon and Pied (1969)
Indonesia equation	$\frac{1}{\rho_t}=\frac{S_w^2}{F\rho_w}+2\left[\sqrt{\frac{V_{sh}^{(2-V_{sh})}}{F\rho_w\rho_{sh}}}\right]S_w^2+\frac{V_{sh}^{(2-V_{sh})}S_w^2}{\rho_{sh}}$	$\sigma_t=\frac{S_w^2}{F}\sigma_w+2\left[\sqrt{\frac{\sigma_w\sigma_{sh}V_{sh}^{(2-V_{sh})}}{F}}\right]S_w^2+\sigma_{sh}V_{sh}^{(2-V_{sh})}S_w^2$	Poupon and Leveaux (1971)
Schlumberger	$\frac{1}{\rho_t}=\frac{S_w^2}{F(1-V_{sh})\rho_w}+\frac{V_{sh}S_w}{\rho_{sh}}$	$\sigma_t=\frac{S_w^2}{F(1-V_{sh})}\sigma_w+V_{sh}\sigma_{sh}S_w$	Schlumberger (1972)
Dual Water Equation	$\frac{1}{\rho_t}=\frac{S_w^2}{F_0\rho_w}+\frac{\left(\frac{1}{\rho_{bw}}-\frac{1}{\rho_f}\right)V_0Q_vS_w}{F_0}$	$\sigma_t=\frac{S_w^2}{F_0}\sigma_w+\frac{(\sigma_{bw}-\sigma_w)V_0Q_vS_w}{F_0}$	Clavier et al. (1977)
Juhasz	$\frac{1}{\rho_t}=\frac{S_w^2}{F\rho_w}+\left(\frac{1}{F_{sh}\rho_{sh}}-\frac{1}{\rho_w}\right)\frac{V_{sh}\phi_{sh}S_w}{\phi}$	$\sigma_t=\frac{S_w^2}{F}\sigma_w+\left(\frac{\sigma_{sh}}{F_{sh}}-\sigma_w\right)\frac{V_{sh}\phi_{sh}S_w}{\phi}$	Juhasz (1981)

ρ_t is the total resistivity of the material (Ωm), ρ_f is the resistivity of the aqueous pore fluid (Ωm), ρ_{sh} is the resistivity of the shale (clay) (Ωm), ρ_{bw} is the resistivity of the bound water in the dual-water model (Ωm), the analogous symbols with σ replacing ρ refer to the equivalent conductivities, S_w is the saturation of the aqueous fluid (dimensionless), V_{sh} is the volume fraction of shale (clay) in the material (dimensionless), V_0 is the volume of shale associated with one coulomb of excess charge ($m^3\ C^{-1}$), ε is an ad hoc constant in the original model (dimensionless), F is the formation factor (dimensionless), F^* is the shaley sand formation factor (dimensionless), F_0 is the formation factor associated with the total porosity (dimensionless), Q_v is the excess of surface charge due to adsorption of ions within the electrical double or triple layer normalized by pore volume ($C\ m^{-3}$), B is ionic surface mobility ($m^2\ s^{-1}\ V^{-1}$), ϕ is the porosity of the sample (dimensionless), and ϕ_{sh} is that of the shale component (dimensionless).

(1960) model is the exception, associating a smaller connectedness to the shale volume as indicated by the exponent 2.

The Waxman and Smits (1968) and dual-water (Clavier et al., 1977) models are based on the ionic double layer observed in shaley sandstones. In reality, the conductivity of the shale component is a function of the CECs of the various types and abundances of clay minerals that are present. Since the cations are exchanged primarily at broken bonds on mineral surfaces, the phenomenon tends to be surface area-dependent rather than controlled simply by the volume of clay minerals. This implies that fine-grained clay has a higher exchange capacity than a coarse-grained form of the same clay volume, and this observation is confirmed by experimental data. In fact, surface conduction studies can now benefit from electrochemical modeling of the electrical double or triple layer structures that form at mineral/fluid interfaces (see **Section 11.04.4**) and it turns out that clays have higher-surface-conductivity contributions than a silica matrix not because their surfaces are intrinsically more conductive, but because they have large internal surface areas (Revil and Leroy, 2001).

The problem with the Waxman and Smits model is that it contains parameters that cannot be obtained from borehole measurements, which give only the approximate volume of shale and the total porosity of the rock. Neither is there any estimation of grain or pore size from downhole measurements. The CEC can be measured in the laboratory (see Ross and Ketterings, 2011 for a summary of the different methods), albeit on crushed rock samples that may or may not represent the *in situ* reality. Consequently, the model equations that use cation-exchange data have been modified to variants that substitute quantities that can be measured on logs as surrogate variables. The most widely known of these is the dual-water model introduced by Clavier et al. (1977), which, when generalized, takes the form

$$S_{wt}^n-\left[S_b\left(1-\frac{a\rho_w}{\rho_b}\right)\right]S_{wt}-\frac{a\rho_w\phi_t^m}{\rho_t}=0 \qquad [45]$$

where $\rho_b=\rho_{sh}\phi_{tsh}^2$, $S_b=V_{sh}\phi_{tsh}/\phi_t$, and $\phi_t=\phi_{eff}+V_{sh}\phi_{tsh}$ and where ρ_b is bound-water resistivity, S_b is bound-water saturation, S_{wt} is total water saturation (including S_b), ϕ_{tsh} is shale porosity (some weighted average of the neutron and density porosities of shale), ϕ_{eff} is effective porosity, ϕ_t is total porosity (including the shale porosity term), and a is Winsauer et al.'s (1952) modification to Archie's law.

Although the Simandoux and dual-water equations were developed from different models, they do have a number of features in common (and with many other shaley sandstone equations):

1. At zero shale content, the equations collapse to the combined Archie's equation (although often with the saturation exponent n set to 2).
2. The equation forms can be written as polynomials and solved reiteratively.

3. The input shale properties are typically drawn from shales between the reservoir units, which may not be representative of clays within the reservoir zones.

11.04.4 Surface Conduction

11.04.4.1 The Electrical Double Layer and Surface Conduction

11.04.4.1.1 The electrical double layer (EDL)

When a freshly broken mineral surface is exposed to an aqueous pore fluid, it attracts and adsorbs positive and negative ions from the fluid (Adamson, 1976; Davis and Kent;, 1990; Dukhin and Derjaguin, 1974; Glover et al., 1994; Hunter, 1981; Sposito, 1989). In the following discussion, we consider a silica surface. However, the same approach may be taken for any mineral, for example, Revil and Leroy (2001) who applied a triple-layer model to clay minerals.

A quartz surface, such as that shown in green in Figure 11, provides two types of neutral surface group on the silica surface, a doubly coordinated siloxal $>Si_2O^o$ group, which can be considered inert and is not shown in the figure, and a singly coordinated silanol $>SiOH^o$ group (where $>$ represents the matrix) (e.g., Iler, 1979; Revil and Glover, 1997, 1998; Revil et al., 1999a). The silanol group reacts readily to give $>SiOH_2^+$ when $pH < pH_{pzc}$ and $>SiO^-$ when $pH > pH_{pzc}$. Both of these are shown in the inner Stern layer in Figure 11, noting that the negative group is surrounded by a hydration shell. The pH_{pzc} is the pH at which the surface has no net charge (i.e., $[>SiOH_2^+] = [>SiO^-]$, where the square brackets indicate concentration) and is called the point of zero charge. For quartz, $pH_{pzc} \approx 3$ (Lorne et al., 1999a,b).

These surface sites, which form the inner Stern layer, react with ionic species in the bulk fluid. If we consider a 1:1 electrolyte with single-valency anions and cations, such as NaCl,

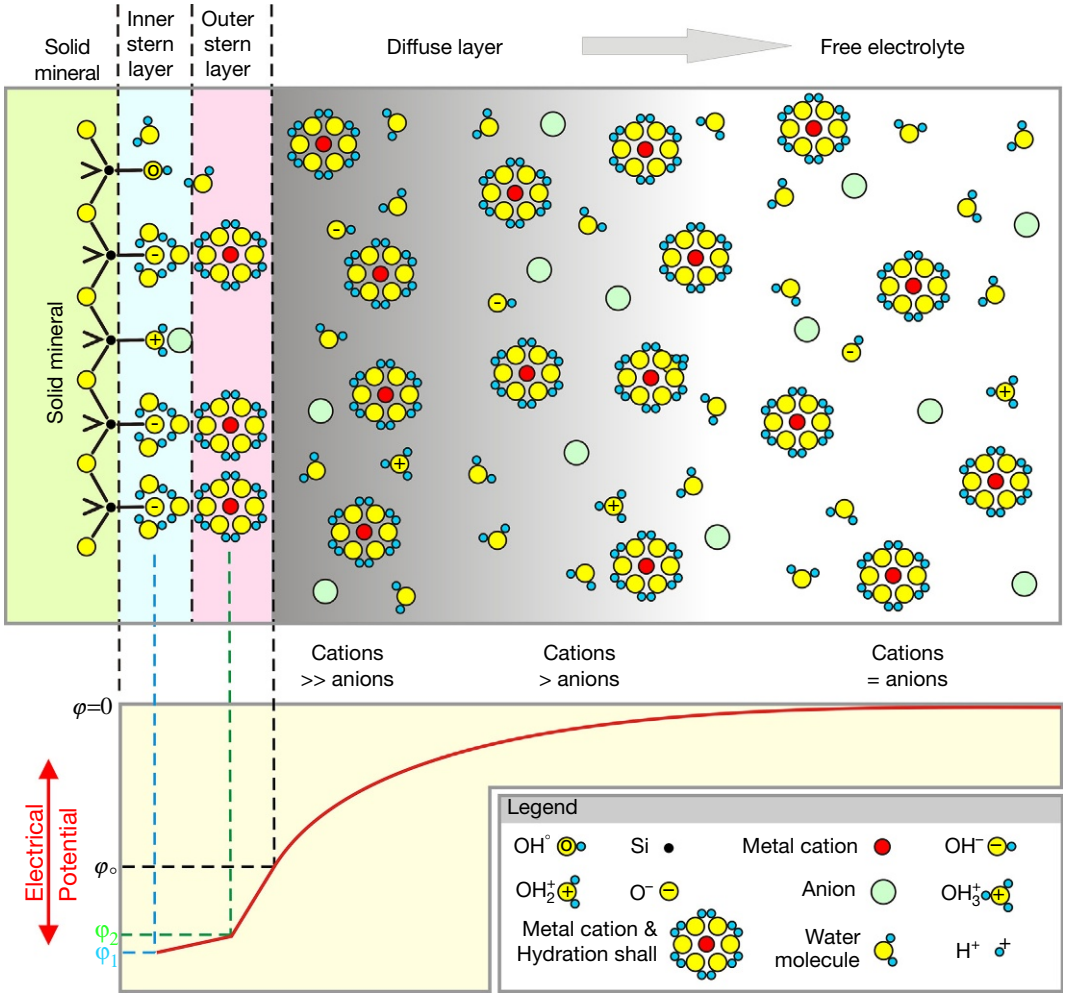


Figure 11 A schematic representation of the electrical double layer that exists at the interface between the rock matrix and the pore water. The rock is shown on the left-hand side in green. Here, it is silica, which forms three types of surface site (positive $SiOH_2^+$, neutral SiOH, and negative SiO^-). The relative concentrations of these depend upon the fluid pH. At reservoir fluid pH, the negative sites dominate. Hydrated positive ions (in red here) are adsorbed to these sites. The adsorbed positive ions cannot cancel out all of the negative surface charge, so there exists a fluid diffuse layer that has an excess positive charge (more hydrated cations than anions), which falls off exponentially until there are equal numbers of positive and negative ions (the bulk fluid). The thickness of the diffuse layer depends upon fluid concentration. At low concentrations, it may be several microns thick, which is sufficient to ensure that a rock with small pores has all of its aqueous fluid in the form of the diffuse layer rather than as bulk fluid. In this case, there is no bulk fluid conduction, only surface conduction mediated through the EDL. Modified from Glover PWJ and Jackson MD (2010) Borehole electro-kinetics. *The Leading Edge* 29: 724–728, http://dx.doi.org/10.1190/1.3447786.

and we consider a geologically reasonable pore fluid pH range of 6–8, the surface reactions can be written as (Davis et al., 1978; Revil and Glover, 1997; Revil et al., 1999a,b)

$$>\mathrm{SiOH^o} \overset{K_{(-)}}{\Leftrightarrow} >\mathrm{SiO^-} + \mathrm{H^+} \quad [46]$$

and

$$>\mathrm{SiOH^o} + \mathrm{Me^+} \overset{K_{Me}}{\Leftrightarrow} >\mathrm{SiOMe^o} + \mathrm{H^+} \quad [47]$$

where the metal cation is Me^+, for example, Na^+, K^+, or some other single-valence cation. The positive surface site $>SiOH_2^+$ does not occur in eqns [46] and [47] because at $pH > 6$, $[>SiOH_2^+] \approx 0$. Hence, there result three types of site, two neutral ones ($>SiOH^o$ and $>SiOMe^o$) and one negative ($>SiO^-$) (Revil and Glover, 1997; Revil et al., 1999a,b). The metal ions that take part in these surface adsorption reactions form the outer Stern layer, as shown in Figure 11.

The position of the equilibrium of eqn [46] is given by the disassociation constant for dehydrogenization of silanol surface sites $K_{(-)}$. If $K_{(-)}$ is large, eqn [46] has an equilibrium towards the right, giving very many more $>SiO^-$ sites than $>SiOH^o$ sites, and vice versa. The position of the equilibrium in eqn [47] is given by the binding constant for cation (sodium) adsorption on quartz K_{Me}. If K_{Me} is large, eqn [47] has an equilibrium towards the right, where there are very many more $>SiOMe^o$ sites than $>SiOH^o$ sites, and vice versa. So $K_{(-)}$ and K_{Me} describe the relative concentrations of the three surface sites. They are unitless.

The final parameter of interest is the total surface site density Γ_s^o (in m^{-2}), which is the sum of the surface site densities for each of the three types of surface site. For quartz and most geologic pore fluids (where $pH \gg pH_{pzc}$). $K_{(-)}$ is large and the inner Stern layer has an overall negative charge, which attracts cations from the bulk fluid. This is a self-limiting process because the number of negative sites is limited. However, the bulk fluid becomes depleted to some extent in cations. This depleted zone is called the diffuse layer as shown in Figure 11. It should be noted that the ions in the outer Stern layer are not usually considered to be mobile, while the ions that make up the diffuse layer are mobile.

The adsorbed ions of the outer Stern layer are not sufficient to balance the negative surface charge associated with inner Stern layer. Hence, the net charge of the two Stern layers remains slightly negative as shown by the bottom portion of Figure 11. It is this negative charge that attracts cations from the bulk solution and ultimately generates a diffuse layer that is depleted in cations.

The global requirement for electrical neutrality ensures that the depletion is greatest near the Stern plane and reduces exponentially with distance from the Stern plane until the diffuse layer blends seamlessly with the bulk fluid. During this process, the electric potential reduces exponentially until it reaches zero in the bulk fluid as shown in the bottom portion of Figure 11.

11.04.4.1.2 Thickness of the electrical double layer

Since the diffuse layer has no well-defined boundary towards the bulk fluid, its width is difficult to define. The traditional measure of the width of the diffuse layer is the Debye screening length that is given by

$$\chi_d = \sqrt{\frac{\varepsilon_o \varepsilon_r k_b T}{2000 N e^2 I_f}} \quad [48]$$

where $\varepsilon_o = 8.854 \times 10^{-12}\ \mathrm{F\,m^{-1}}$ is the electric permittivity *in vacuo* of the fluid (Lide, 2012), ε_r is the relative electric permittivity of the fluid (no units), T is the temperature (in K), and I_f is the ionic strength, which is given by

$$I_f = \frac{1}{2}\sum_i^n Z_i^2 C_i^f \quad [49]$$

where there are i ionic species in solution, Z_i is the valency of each ionic species, and C_i^f is the concentration of each ionic species in solution. The factor of 2000 arises due to the units for ionic strength here being $\mathrm{mol\,l^{-1}}$. This equation is often found cited with a 2 in place of the 2000, in which case the units for I_f would be $\mathrm{mol\,m^{-3}}$.

In most geologic fluids in the upper crust the ionic contribution to fluid conductivity from acid ions $[H_3O^+]$ and alkali ions $[OH^-]$ are much smaller than the contribution from the electrolyte, for example, $[Na^+]$ and $[Cl^-]$ and for an NaCl electrolyte. Hence, for the most common scenarios where $C_f > 10^{-5}\ \mathrm{mol\,l^{-1}}$ and $5 < pH < 9$, we can say that $I_f \approx C_f$. In fact, recent experimental studies have shown that it is extremely difficult to have a pore fluid with $C_f < 10^{-5}\ \mathrm{mol\,l^{-1}}$ in full chemical equilibrium with a sandstone (Walker et al., 2014) because the salinity of the fluid is raised by dissolution. It is likely that this observation is also true for many other rock types, making the approximation $I_f \approx C_f$ valid in most geologic scenarios.

The most common measure of the effective width of the diffuse layer is twice the Debye screening length (e.g., Revil and Glover, 1997). Figure 12 shows how the Debye screening length and the effective width of the diffuse layer vary with the salinity of the bulk fluid. The diffuse layer is thicker at low salinities than at high salinities. This is simple to explain physically: A low-salinity bulk fluid has fewer cations per unit volume than a high-salinity fluid. Hence, it will take a greater volume of low-salinity fluid to provide cations to adsorb to the surface and to completely saturate all the surface sites than would be the case for a high-salinity fluid with more cations per unit volume. Since the area of the rock–fluid interface is fixed, the diffuse layer must be thicker for low-salinity fluids.

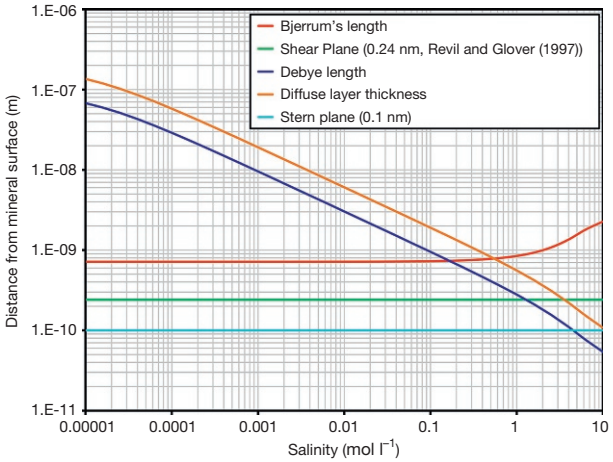


Figure 12 Characteristic length scales related to the electrical double layer in saturated porous media.

For low-salinity fluids and rocks with small pore sizes, it is possible for the width of two diffuse layers (one associated with each of two opposing grains of the rock) to be larger than the distance between the grains. In this case, there is no bulk fluid and all of the fluid between the two grains is composed of the diffuse layers of each EDL. Since the diffuse layer is charged (positively for silica geologic pHs), ions of the same charge can be excluded from entering these pores and pore throats, with the overlapping diffuse layers acting as a charge selective barrier. Such a scenario can occur in clay-rich rocks or clay concentrations within heterogeneous reservoir rocks.

For all salinities, the thickness of the diffuse layer is significantly larger than the thickness of the Stern layer (the blue line in Figure 12), which may be taken as being of the order of the diameter of a hydrated metal ion (viz., 10^{-10} m).

Figure 12 also shows the distance of the shear plane from the Stern plane $\chi_\zeta = 2.4 \times 10^{-10}$ m (the green line in Figure 12), which was calculated by Revil and Glover (1997) and the Bjerrum length (the red line in Figure 12). The Bjerrum length λ_b is the distance at which the electrostatic interaction between the two elementary charges is comparable in magnitude to the thermal energy scale. It is given by $\lambda_b = e^2/4\pi\varepsilon_r\varepsilon_o k_b T$. For an aqueous solution of NaCl, we find that the Bjerrum length remains constant at $\lambda_b = 7.16 \times 10^{-10}$ m for salinities less than about 0.1 mol l^{-1}, increasing substantially as salinity increases past 0.1 mol l^{-1} until it reaches a value of 13.5×10^{-10} m at 3.98 mol l^{-1}.

At length scales larger than λ_b, fluid ions are more affected by thermal agitation than the presence of the Stern layer. At length scales smaller than λ_b, fluid ions are significantly affected by the presence of the Stern layer. Indeed, the local field strength is sufficient to alter the viscosity and the permittivity of the fluid considerably. It should be noted that the Bjerrum length is greater than the EDL thickness for all salinities greater than about 0.55 mol l^{-1} (dashed line in Figure 12). In other words, interionic electrostatic interactions are significant for the entire EDL, including all of its mobile fraction ($\chi > \chi_\zeta = 2.4 \times 10^{-10}$ m) for salinities greater than 0.55 mol l^{-1}, and it is these electrostatic interactions that may increase the viscosity of the fluid in the double layer and decrease the local permittivity.

11.04.4.2 Basic Theory of Surface Conduction

Originally, it was thought that surface conduction was only significantly developed in clays where the presence of Al^{3+} ions in place of Si^{4+} ions led to a charge deficit that was counterbalanced by the adsorption of cations from solution that then took part in surface conduction. The description in the previous section updates this view by making the process general to all minerals. Further development stemmed from the approach taken by Waxman and Smits (1968), whose model has already been mentioned in Sections 11.04.2.4 and 11.04.3.5.

11.04.4.2.1 Waxman and Smits model

It became increasingly clear throughout the 1950s and 1960s that Archie's simple empirical laws could not be applied to rocks that contained substantial clay fractions because of the significant surface conduction they exhibited. One of the first, empirical, and most successful formulations to overcome this problem was proposed by Waxman and Smits (1968), who included a surface conduction term in Archie's equation:

$$\sigma_{\text{eff}} = \frac{1}{F^*}(\sigma_f + BQ_v) \qquad [50]$$

The extra surface conduction is accounted for by the term BQ_v, where B (in Sm^2 mequiv.$^{-1}$) represents the mobility of surface cations and Q_v (in mequiv. m^{-3}) is the volume concentration of clay exchange cations (mequiv. = milliequivalents). Both parameters are usually obtained by fitting the model to experimental data, but the CEC is sometimes measured independently. It is worth noting that the single value for the Waxman and Smits formation factor F^* shows that the model assumes that the pathways for conduction through the bulk fluid and via surface conductivity are the same.

The critical point in the Waxman and Smits model is that the surface conduction term is composed of the product of an element that describes the number density of charge carriers on the mineral surface, which is represented by Q_v, and an element describing the mobility of those ions, which is represented by B.

11.04.4.2.2 Ionic surface electrical conduction models

In 1994, Glover et al. (1994) recognized that a model for surface conduction would have to describe the theoretical framework behind both the number density and the mobility of surface charge carriers. Restricting themselves to the first of these challenges Glover et al. (1994) attempted to create a model to calculate the concentration of surface charge carriers, which they called ISCOM1 (ionic surface concentration model). This basic model was quickly superseded by a full model for surface conduction that grew from ISCOM1 but that took account of the full expression of the electrical double layer and included surface mobilities (Revil and Glover, 1997, 1998; Revil et al., 1999a,b). Various versions of the latter model now exist and are widely applied for expressing the surface conduction in porous media (e.g., Revil and Leroy, 2001).

11.04.4.3 Application to Clean Rocks

An application of the Revil and Glover (1997) model to clean quartz-dominated rocks can be found in the last part of that paper. The equations are developed for a quartz matrix immersed in an electrolyte. The application of the Revil and Glover approach allows the calculation of the Stern plane potential (Figure 13) and hence the surface charge density as a function of pH and electrolyte salinity (Figure 14).

The use of these parameters allows the calculation of the electrical conductivity of the rock as a function of salinity and pH once certain microstructurally defined parameters are known (Figure 15). These parameters are the formation factors related to electrical transport through the bulk pore fluid and via surface conduction and their respective length scales. Figure 15 shows a comparison of the model for all the surface sites on the quartz being negative (the situation in the previous theory by Glover et al., 1994), which is shown by the dotted line, and the more sophisticated Revil and Glover (1997) model. The latter model is now used more commonly.

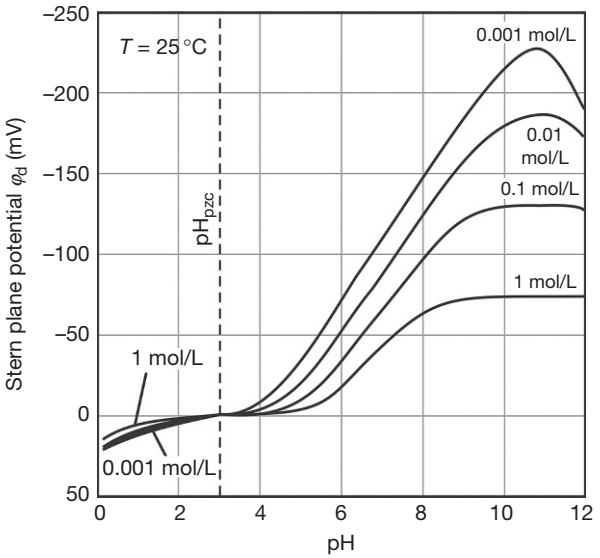


Figure 13 Stern plane potential for a quartz rock saturated with a monovalent electrolyte calculated from the Revil and Glover (1997) theory as a function of pH and electrolyte salinities at $T = 25$ °C. The dashed line marked 'pzc' is the point of zero charge; it corresponds to that pH at which there are equal densities of positive and negative sites on the quartz surface. When this is the case, the surface is neutral and no electrical double (or triple) layer is developed leading to no surface conduction.

11.04.4.4 Application to Clay-Rich Rocks

The electrical conductivity of clays is a far more difficult problem due to the variability of both the composition and the structure of the clay minerals. Initially, the problem was approached using a macroscopic model (Revil and Glover, 1998; Revil et al., 1998), but subsequent works by Revil and Leroy (e.g., Leroy and Revil, 2004; Revil and Leroy, 2001, 2004) have led to a full theoretical understanding of the electrochemical properties of clay minerals and ionic transport in porous shales.

11.04.5 Frequency-Dependent Electrical Properties

So far, we have only considered steady-state electrical flow, that is, when a constant potential difference is applied to a material and a constant electric current flows. Often, the applied potential will vary with time, either as a transient or with a regular repeating manner such as a harmonic pattern. This section considers such time-dependent behavior. The theoretical description of such behavior needs to take account of the fact that current flow may not keep in step with the applied potential difference due to the way that a material may store charge over short timescales or the fact that some electrochemical processes take time to occur.

There are two processes, (i) the transport of charge, which is described by the electrical conductivity of a material, and (ii) charge storage, which is described by the dielectric polarization of a material.

These two process are described fundamentally by

and $$\vec{J} = \sigma \vec{E} \qquad [51]$$

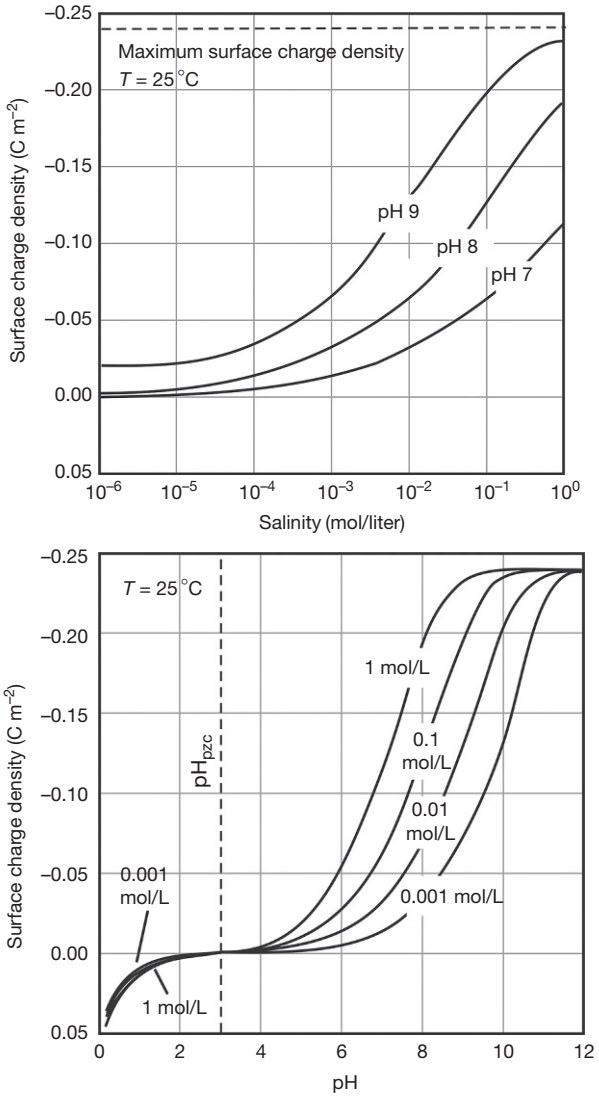


Figure 14 The surface charge density for a quartz rock saturated with a monovalent electrolyte calculated from the Revil and Glover (1997) theory as a function of salinity and pH at $T = 25$ °C. The dashed line marked 'pzc' is the point of zero charge; it corresponds to that pH at which there are equal densities of positive and negative sites on the quartz surface. When this is the case, the surface is neutral and no electrical double (or triple) layer is developed leading to no surface conduction.

$$\vec{D} = \varepsilon \vec{E} \qquad [52]$$

where $\vec{E}$ is the electric field strength (in V m^{-1}), $\vec{J}$ is the current density (in A m^{-2}), $\vec{D}$ is the electrical displacement (in C m^{-2}), σ is the electrical conductivity (in S m^{-1}), and ε is the dielectric permittivity (in F m^{-1}), which is often written as the product of the permittivity of free space and the dimensionless relative permittivity or dielectric constant of the material ε_r; hence, $\varepsilon = \varepsilon_o \varepsilon_r$. The electric field strength, current density, and electrical displacement are vectors, while the electrical conductivity and dielectric permittivity are frequency-dependent tensors that take account of the anisotropy of the material.

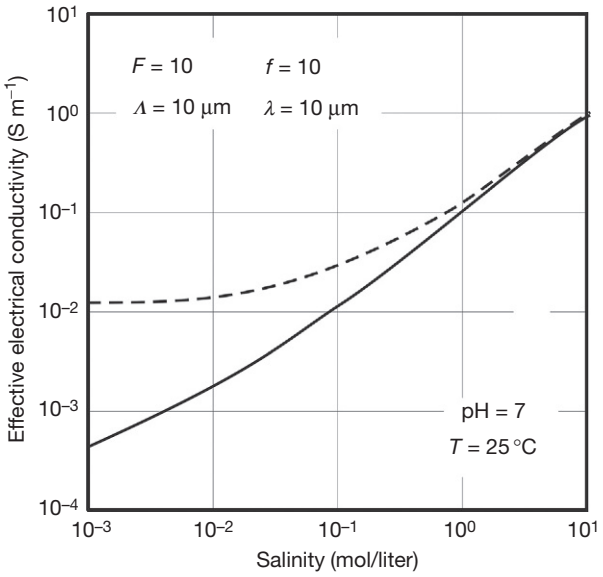


Figure 15 The effective electrical conductivity for a quartz rock saturated with a monovalent electrolyte at pH = 7 and T = 25 °C as a function of the electrolyte conductivity calculated from the Revil and Glover (1997) theory. The dotted line is the relationship with a quartz surface saturated with negative sites as in the Glover et al. (1994) model, and the solid line is the Revil and Glover (1997) theory. Symbols F and f are the formation factors that relate to conduction through the bulk fluid and associated with surface conduction, respectively (cf. eqn [38]). Λ and λ are the characteristic length scales associated with conduction through the bulk fluid and associated with surface conduction, respectively (cf. eqn [39]). Reproduced from Revil A and Glover PWJ (1997) Theory of ionic-surface electrical conduction in porous media. *Physical Review B - Condensed Matter and Materials Physics* 55(3): 1757–1773.

11.04.5.1 Complex Conduction

The conductive aspect of charge movement, described by [51], can be described using a complex conductivity (Guéguen and Palciauskas, 1994):

$$\tilde{\sigma} = \sigma' + i\sigma'' \qquad [53]$$

where $i = \sqrt{-1}$. We will use the convention that all complex quantities have a superscripted tilde. Hence, the complex conductivity is equal to the sum of the real (σ') part and the imaginary ($i\sigma''$) part.

Since the electrical resistivity is the inverse of the conductivity, we may also write (Guéguen and Palciauskas, 1994)

$$\tilde{\rho} = \tilde{\sigma}^{-1} = \rho' - i\rho'' \qquad [54]$$

The real part of the complex conductivity describes energy transport, while the imaginary part describes energy dissipation.

11.04.5.2 Dielectric Permittivity and Polarization

When an electric potential difference is applied to a geomaterial, charges move in response to the electric field. Mobile charges form an electric current that is described by the conductivity of the material, while bound charges also move within their ability to give a polarization of charges.

For example, the electrons that surround the atom of a molecule redistribute themselves with respect to an applied electric field, and the electron 'cloud' becomes distorted towards the positive of the applied external field. This is called electronic polarization. From a distance, the atom has been polarized with a slight excess of negative charge facing the positive of the external field and a slight excess of positive charge facing the negative pole of the applied field. Electrons have a small mass, and hence, the polarization occurs very quickly (of the order of 10^{-15} s). This is called the 'relaxation time' of the process. The relaxation time associated with electron polarization is the same order as the period of visible light. It is for this reason that the electronic contribution to permittivity is related to the refractive index, n, of a material ($\varepsilon_r = 1/n^2$). Hence, the permittivity from this source for most solid materials is in the range 1–4.

There are other sources of polarization and each operates at a different timescale. The next process, in order of increasing timescale, is atomic or ionic polarization, where atoms or ions in crystalline and amorphous solids polarize with respect to an applied electric field. As ions are heavier than electrons, the relaxation time for this process is longer (10^{-12}–10^{-13} s), but still very fast. The permittivity from this source is in the range 1–11. At microwave frequencies and below (i.e., less than 10^{12} Hz), we can say that most solid materials, including all minerals, have a relative permittivity of about 4–15 because relative permittivities from different sources add algebraically.

Dipolar polarization is another process, this time affecting materials that already have polar molecules. These polar molecules align themselves with an applied electric field. Once again, the timescale is slower (10^{-8} to 10^{-12} s), which reflects the masses of the polar molecules and the ease with which they may rotate. These polarizations can be large (5–100). The relative permittivity of water (~80) is due to electronic polarization and dipolar polarization, of which the latter is by far the largest contributor, while that of most oils is about 2 because oils contain almost no polar substances and the permittivity is controlled by electronic polarization.

Heterogeneous materials, which include all rocks, have an additional mechanism that is associated with mobile charge carriers that become stuck at or accumulate at interfaces. This occurs in porous and cracked rocks where ions in the pore fluid respond to the application of an external electric field, but solid mineral grains get in the way. Ions accumulate near the grain surface and their charge counteracts further motion of ions. The process is sometimes called the Maxwell–Wagner process, but this term was originally coined for a material consisting of one material of one relative permittivity dispersed in another of a second relative permittivity. It is also called interfacial polarization. Interfacial polarization has a relaxation time of 10^{-3} to 10^{-6} s and can be very large. Values of the order of 1000 are not unusual, which is much larger than the relative permittivity of water (~80) or mineral grains (~4–15). Hence, we can say that, as far as relative permittivity is concerned, the heterogeneity and arrangement of the materials that make up a geomaterial are much more important than the relative permittivities of its individual components. Recently, there has been an upsurge in work on dispersion mechanisms in porous media. Particular attention is now being paid to polarization processes occurring in the Stern layer of the electrical double layer or the electrical triple layer (Revil and Florsch, 2010; Revil and Skold, 2011; Schmutz et al., 2010; Skold et al., 2011).

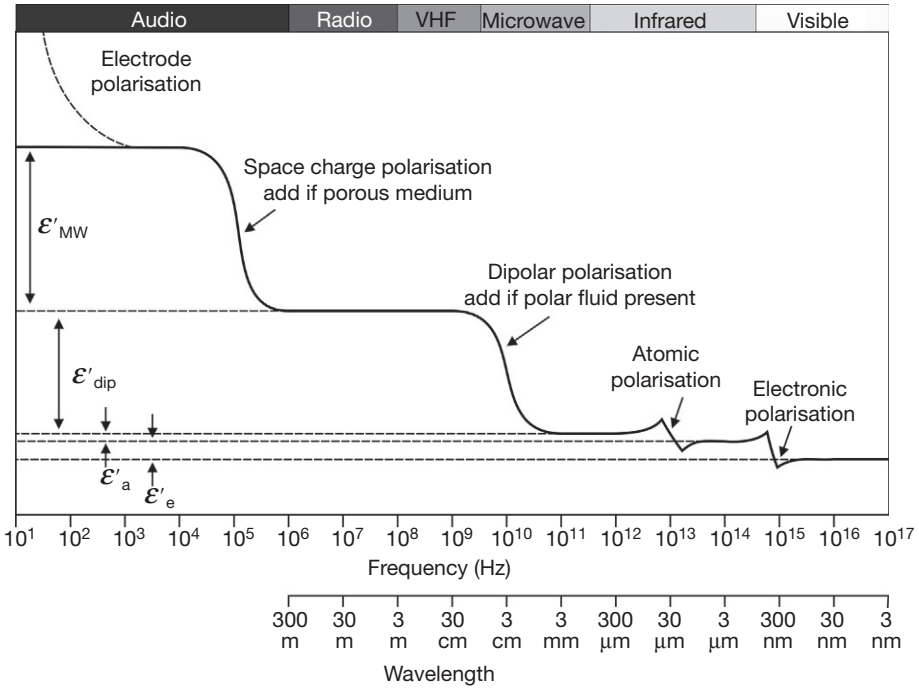


Figure 16 Dielectric polarization processes (e.g., Izgorodina et al., 2009). The dashed line shows where electrode polarization would manifest itself. This is a measurement artifact and hence shown with a dashed line.

Electron, atomic, dipolar, and interfacial dispersion mechanisms are shown in the classic diagram (Figure 16). Here, the relative size and frequency of each process are shown (e.g., Izgorodina, et al., 2009). At frequencies above about 10^{15} Hz, only electronic polarization (ε'_e) contributes to the permittivity. As frequency reduces to about 10^{12} Hz, atomic polarization (ε'_a) also contributes and is added to the electronic contribution. Further decrease in frequency to radiofrequencies sees dipolar contributions (ε'_{dip}) added to the permittivity, and finally, at low frequencies, we see the addition of interfacial polarization (ε'_{MW}). This diagram implies that the permittivity that one must use in any application should be chosen with care, taking into account the frequency range in which one works.

Figure 16 also shows a fifth process called electrode polarization. This is not a property of the material itself, but represents a technical difficulty in making measurements of the electrical properties of geomaterials. It is shown in the figure as a broken line to indicate that it is an artifact and to show where it would manifest itself if it were present in measurements. It arises from the electrochemical reactions that occur at the surface of the electrodes that transport current into and out of the sample rather than any process in the sample itself. In an electrical measurement on a rock the current is carried by a flow of electrons in the wires of the measurement device and the electrodes, but is carried by ions in the electrolyte which occupies the pores. The transfer of charge occurs at the electrode surface through equations such as

$$e^- + Na^+ \leftrightarrow Na \qquad [55]$$

Here, the electron provided by the electrode reacts with the sodium ion from the solution that has been attracted to the electrode and neutralizes it, thus facilitating electrical flow in the pore fluid. The trouble is that these processes take time to occur causing charge to build up at the electrode surface which repels further charges and reduces current flow. The relaxation time is large compared to the processes mentioned previously. Hence, electrode polarization can be a significant problem for measurements made at frequencies lower than about 200 Hz.

There is now growing evidence that other processes exist at lower frequencies still. These are not shown in Figure 16. Electrochemical polarization can occur in the Stern layer of the electrical double or triple layer that coats the grain surfaces. Leroy et al. (2008) measured a dispersion associated with particle size occurring at about 75 mHz for a random pack of spherical grains of modal grain size 187.5 µm and another dispersion that they associated with the characteristic size of asperities on the surface of their grains (i.e., surface roughness) at about 20 Hz, to which they attributed a characteristic scale of 12 µm. A modeling study by Volkmann and Klitzsch (2010) also showed the presence of a low-frequency dispersion, but at slightly higher frequencies.

The existence of these low-frequency polarization mechanisms shows that it is extremely important to control electrode polarization. Section 11.04.5.7 discusses several methods for reducing or removing electrode polarization from measurements.

11.04.5.3 Dielectric Permittivity Models

Figure 16 only shows the real part of the permittivity. In reality, the permittivity is a complex quantity in the same way as the conductivity and resistivity. It can be written as

$$\tilde{\varepsilon} = \varepsilon' - i\varepsilon'' \qquad [56]$$

The classical model for describing the transition of permittivity due to each of the polarization processes is that of the Nobel laureate Peter Debye (Debye, 1929), who developed a model for dipolar liquids and a single relaxation time (i.e., one polarization process with a well-defined timescale):

$$\tilde{\varepsilon}(\omega) = \varepsilon_\infty + \frac{\varepsilon_o - \varepsilon_\infty}{1 + i\omega\tau} \quad [57]$$

where ω is the angular frequency ($=2\pi f$ in rad s^{-1}), τ is the relaxation time (s), and ε_o and ε_∞ are the values of permittivity at low and high frequencies with respect to the critical frequency ($\omega_{crit}=\tau^{-1}$). While the Debye model was developed for dipolar polarization, it has also been applied to other polarization mechanisms. Equation [57] can be separated into its real and imaginary parts (Guéguen and Palciauskas, 1994):

$$\varepsilon'(\omega) = \varepsilon_\infty + \frac{\varepsilon_o - \varepsilon_\infty}{1 + \omega^2\tau^2} \quad [58]$$

and

$$\varepsilon''(\omega) = \frac{(\varepsilon_o - \varepsilon_\infty)\omega\tau}{1 + \omega^2\tau^2} \quad [59]$$

The real and imaginary parts of permittivity according to the Debye model are shown in Figure 17. The solid blue line in the upper part of Figure 17 shows how the normalized in-phase permittivity varies between 0 and 1. The imaginary part of the permittivity (Shown by the solid blue line in the middle part of Figure 17), which describes energy loss, peaks at the critical frequency ($\omega_{crit}=1/\tau$), and almost all of the change occurs within one order of magnitude of the critical frequency. The frequency dependence is also shown in the form of an argand diagram (Bottom part of Figure 17), where the out-of-phase (imaginary) relative permittivity is plotted as a function of the real (in-phase) permittivity. The Debye model (Blue line) plots as a semicircle on an argand diagram, with the center of the circle on the x-axis as shown by the reference mark in the figure.

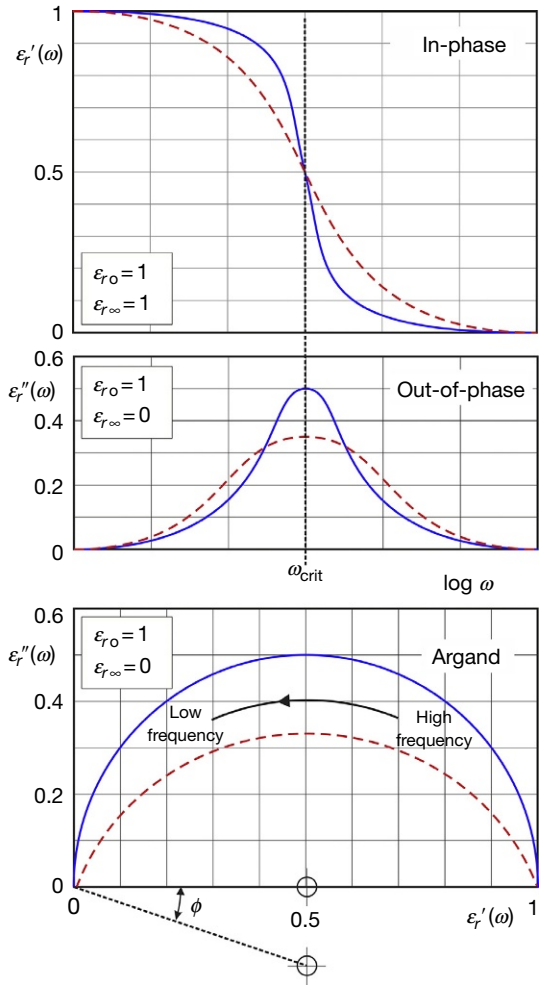


Figure 17 The behavior of the Debye (solid blue) and Cole and Cole (dashed red) dispersion models shown as a function of frequency for each of the in-phase (upper) and out-of-phase (middle)components and as an are Argand diagram (bottom). An arbitrary value of $\alpha = 0.22$ has been used for the Cole and Cole model.

Sometimes, there is difficulty in understanding this complex model from a physical point of view. At frequencies significantly lower than the critical frequency, the polarization occurs quickly enough for it to always be fully developed. Hence, the permittivity is described fully by the real (in-phase) part of the complex permittivity, and the imaginary part is negligible. At frequencies much higher than the critical frequency for the polarization process, there is insufficient time for the polarization to occur at all. Hence, there is no contribution to permittivity from either the in-phase (real) part of the complex permittivity or the imaginary (out-of-phase) part. In between, there exist frequencies that are about the same as the critical frequency. Polarization processes that occurs at these timescales are only partly completed during each cycle of the applied electric potential difference leading to polarization that is out of step with the applied field generating nonzero out-of-phase (imaginary) permittivities, which peak when the applied frequency is the same as the critical frequency.

Although the Debye model has been applied to interfacial polarization, it has been recognized that there is a broader frequency dependence (Chelidze and Guéguen, 1999; Chelidze et al., 1999) for this process. The broader frequency dependence probably arises from there being a spectrum of relaxation times rather than a single relaxation time. It is easy to imagine that a pore structure that is now widely deemed to be fractal (Ruffet et al., 1991a) will have relaxation times associated with interfacial polarization occurring at the different scales of structure exhibited by the pores. A modification of the Debye model by Cole and Cole (1941) takes an ensemble of relaxation times into account:

$$\tilde{\varepsilon}(\omega) = \varepsilon_\infty + \frac{\varepsilon_o - \varepsilon_\infty}{1 + (i\omega\tau)^{1-\alpha}} \quad [60]$$

where α can take values between 0, where it is formally the same as the Debye model, and 1. The Cole and Cole model is shown by the red dashed lines in Figure 17. The angle ϕ (in radians) in Figure 17 and α are related by $\alpha = 2\phi/\pi$ (Glover et al., 1994). The exponent in the Cole and Cole relationship has been related to the fractal dimension of the surface of the grains of the rock by several researchers (see Glover et al., 1994), but none have been sufficiently supported by experimental evidence to be regarded as credible. The fractal dimension, in some of these models, tends towards 2 as the salinity of the pore fluid decreases, which may indicate that the thickness of the EDL has a role to play.

It can be seen in Figure 17 that the in-phase permittivity of the Cole and Cole model varies between the same (normalized) endpoints as in the Debye model, but more gradually. The peak in the out-of-phase permittivity is at the critical frequency, but is broader than the Debye model. The Cole and Cole model also plots as a semicircle on an argand diagram, but now, the center of the semicircle is depressed below the x-axis. The angle ϕ

subtended by the line that joins the center to the orgin ($\varepsilon_r'=0, \varepsilon_r''=0$) is related to the Cole and Cole α parameter. Models that are based on the approach of Dissado and Hill (1984) have been developed that relate the Cole and Cole parameter to the fractal dimension of the grain/fluid interface (Le Méhauté and Crepy, 1983; Ruffet et al., 1991a,b; Wong, 1987).

In 1951, Davidson and Cole (1951) added a third frequency dependence model, which is usually known as the Cole–Davidson model. The Cole–Davidson model is given by

$$\tilde{\varepsilon}(\omega) = \varepsilon_\infty + \frac{\varepsilon_o - \varepsilon_\infty}{(1+i\omega\tau)^\beta} \quad [61]$$

where the parameter β varies between 0 and 1. It is rarely fitted to complex electrical rock data because it is asymmetrical, the out-of-phase component falling off more slowly at frequencies above the critical frequency (Lockner and Byerlee, 1985). However, there are some indications that it may be of use in the modeling of frequency-dependent electrokinetic measurements (see Section 11.04.6.9).

11.04.5.4 Integrated Conduction and Permittivity

Maxwell recognized that there exist not only electric currents which arise from the transport of charges ($\vec{J}_{cond} = \sigma \vec{E}$) but also what he called displacement currents ($\vec{J}_{disp} = \partial \vec{D}/\partial t$), which only exist when the applied electric field varies with time. The total resulting current is the sum of both of these processes:

$$\vec{J}_{total} = \sigma \vec{E} + \frac{\partial \vec{D}}{\partial t} = \sigma \vec{E} + \varepsilon \frac{\partial \vec{E}}{\partial t} \quad [62]$$

If we consider a harmonically varying field (i.e., $E = E_o e^{i\omega t}$), eqn [62] can be written as

$$\vec{J}_{total} = (\sigma + i\omega\varepsilon)\vec{E} \quad [63]$$

Then, we can define

$$\vec{J}_{total} \equiv \sigma_{total}\vec{E} \equiv \varepsilon_{total}\frac{\partial \vec{E}}{\partial t} \quad [64]$$

where σ_{total} and ε_{total} are the conductivity and permittivity arising from processes involving conduction and displacement currents, and hence,

$$\sigma_{total} = i\omega\varepsilon_{total} \quad [65]$$

$$\sigma_{total} = (\sigma + \omega\varepsilon'') + i\omega\varepsilon' \quad [66]$$

and

$$\varepsilon_{total} = \varepsilon' - i\left(\frac{\omega\varepsilon'' + \sigma}{\omega}\right) \quad [67]$$

Hence, the complex electrical properties of the rock can be expressed fully either in permittivity space, in conductivity space, or as resistivity, depending upon which is easiest in the given application.

In each of [66] and [67], the term in parentheses describes the dissipation of energy by conduction currents (σ) and displacement currents ($\omega\varepsilon''$). These equations are shown graphically in Figure 18. Here, it can be seen that the out-of-phase permittivity that describes energy dissipation has two terms: one for the displacement currents that peaks at the critical frequency and is associated with the interplay between

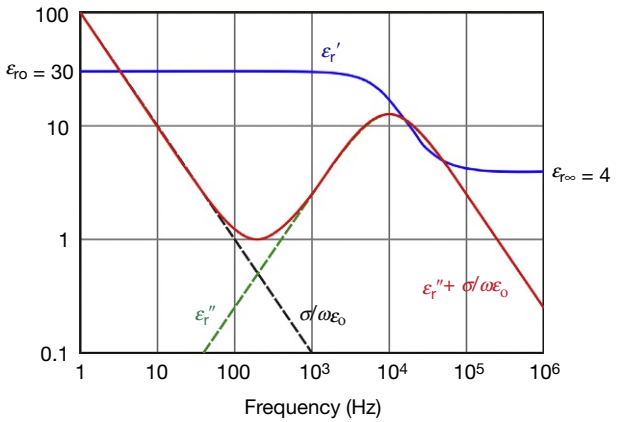


Figure 18 The behavior of the electric permittivity of geomaterials as a function of frequency. The in-phase (real) permittivity is represented by the solid blue curve. The out-of-phase (imaginary) part, which is associated with energy dissipation (solid red curve), is made up of a part related to dissipation due to displacement currents (green dashed line) and due to conduction currents (black dashed lines). In this example, we have taken $\varepsilon_{r\infty} = 4$ as typical as the permittivity for solid minerals and a value of $\varepsilon_{ro} = 30$ to represent the permittivity of a dipolar fluid.

polarization at the critical frequency and the frequency of the applied electric field ε_r'', and one for the conduction currents that increases with decreasing frequency noting that the permittivity is higher for greater salinities ($\sigma/\omega\varepsilon_o$). The contribution from the conductivity of the rock is particularly important. If the salinity is increased, the conductivity of the fluid and of the rock increases, with the $\sigma/\omega\varepsilon_o$ contribution at any given frequency increasing. By contrast, decreasing the saturation of the rock, while keeping the salinity of the saturating fluid constant, decreases the $\sigma/\omega\varepsilon_o$ contribution (Knight et al., 1985).

The energy dissipation is characterized by the ratio of the imaginary part of ε_{total} to its real part or, equivalently, the real part of σ_{total} to its imaginary part, and this ratio is defined as equal to tan δ, where δ is called the loss tangent. Hence,

$$\frac{\sigma + \omega\varepsilon''}{\omega\varepsilon'} \equiv \tan\delta \quad [68]$$

Large values of tan δ indicate high energy loss from the system.

Equation [68] can be written in the form

$$\tan\delta \equiv \frac{\sigma}{\omega\varepsilon'} + \frac{\varepsilon''}{\varepsilon'} \quad [69]$$

where the first term represents conductive energy loss and the second term represents energy loss associated with relaxation processes. At low frequencies, the conductive energy loss becomes significant if $\sigma/\varepsilon' \gg \omega$. For example, for water with a conductivity of 1 S m^{-1} (which is roughly that of a 0.1 mol l^{-1} solution of NaCl at 20 °C) $\sigma/\varepsilon' \approx 1.4 \times 10^9$ rads^{-1}. If the requirement for conductive losses to begin to be significant is that the parameter σ/ε' is just bigger than ω, which would occur as ω dropped lower than 1.4 Grad s^{-1} (220 MHz). However, the requirement is that the parameter σ/ε' is much greater than ω. Let us suppose that the much greater than symbol requires one parameter to be at least two orders of magnitude greater than the other. This would indicate that conductive losses will be significant for frequencies lower than 14 Mrad s^{-1} (2.2 MHz).

Electrical data may also be given as the magnitude of the complex vector together with the phase angle $\varphi = \tan^{-1}(x''/x')$, where x may be conductivity, resistivity, or permittivity.

11.04.5.5 Mixing Models for Complex Electrical Properties

The dielectric response of a geomaterial that is 100% saturated with a fluid can be considered as a two-component mixture with a matrix dielectric constant (relative permittivity) ε_{rs} and a fluid dielectric constant (relative permittivity) ε_{rf}. In general, both are complex and vary with frequency. However, at the frequencies less than 100 GHz, we can consider the dielectric constant of the matrix and that of gases ε_{rg} and oils ε_{ro} to be wholly real and independent of frequency (i.e., $\varepsilon_{rs} = \varepsilon'_{rs}$, $\varepsilon_{rg} = \varepsilon'_{rg}$, and $\varepsilon_{ro} = \varepsilon'_{ro}$), leaving only the complex dielectric constant of water $\varepsilon_{rw}(\omega) = \varepsilon'_{rw}(\omega) + i\varepsilon''_{rw}(\omega)$. At frequencies below the critical frequency for dipolar relaxation (about 10^{10} Hz), it is possible to assume that the imaginary part of the complex dielectric constant of water is negligible ($\varepsilon''_{rw} \approx 0$), which allows us to write

$$\varepsilon_{rw} = \varepsilon'_{rw}(1 - i\tan\delta_w) \cong \varepsilon'_{rw} - i\frac{\sigma_w}{\varepsilon_o\omega} \qquad [70]$$

Hence, from the point of view of its constituents, the dielectric response of the fully water saturated rock can be fully defined by the three parameters, ε'_{rs}, ε'_{rw}, and σ_w. However, from the point of view of the saturated rock as an ensemble, the dielectric constant at low frequencies can also be written as

$$\widetilde{\varepsilon}_{rr} \equiv \varepsilon'_{rr} - i\frac{\sigma_r}{\varepsilon_o\omega} \qquad [71]$$

where ε'_{rr} is the effective dielectric constant of the composite medium and σ_r is the effective electrical conductivity of the composite medium. Since both [70] and [71] must be true for the same composite medium, it can be inferred that ε'_{rr} and σ_r depend on the four parameters, ω, ε'_{rs}, ε'_{rw}, and σ_w.

Scaling arguments (Guéguen and Palciauskas, 1994) imply that the ratio $\widetilde{\varepsilon}_{rr}/\widetilde{\varepsilon}_{rs}$ depends only on the ratio $\widetilde{\varepsilon}_{rw}/\widetilde{\varepsilon}_{rs}$ and the volume fractions of the two phases

$$\frac{\widetilde{\varepsilon}_{rr}}{\widetilde{\varepsilon}_{rs}} = f\left(\frac{\widetilde{\varepsilon}_{rw}}{\widetilde{\varepsilon}_{rs}}, \phi\right) \qquad [72]$$

with

$$\frac{\widetilde{\varepsilon}_{rw}}{\widetilde{\varepsilon}_{rs}} = \frac{\varepsilon'_{rw}}{\varepsilon'_{rs}}(1 - i\tan\delta_w) \quad \text{and} \quad \tan\delta_w = \frac{\sigma_w}{\omega\varepsilon'_{rw}} \qquad [73]$$

The functional form of $f()$ is determined by the rock microstructure.

11.04.5.5.1 The CRIM model

One form of $f()$ was arrived at empirically and is most accurate at very high frequencies (above 100 MHz). This model is called the complete refractive index method (CRIM) and is given by

$$\widetilde{\varepsilon}_r^{1/2} = \widetilde{\varepsilon}_{rw}^{1/2}\,\phi + \widetilde{\varepsilon}_{rs}^{1/2}(1 - \phi) \qquad [74]$$

At such high frequencies, the square root of the dielectric constant is inversely proportional to the phase velocity, which makes [74] a type of time-average equation describing the transmission of an electromagnetic wave through a composite material composed of matrix $(1 - \phi)L$ and pore fluid ϕL, where L is the overall length of the sample and ϕ is the porosity.

An interesting feature of the CRIM model is that in the low-frequency limit, where $\omega \to 0$ and $\delta_w \to \pi/2$, it collapses to give $\sigma = \sigma_w\phi^2$, which is Archie's first law with a cementation exponent of 2.

11.04.5.5.2 The Lichtenecker–Rother equation

A more general mixing law that can be applied to permittivities is the empirical Lichtenecker–Rother equation:

$$\widetilde{\varepsilon}_r = \left(\sum_{i=1}^{n} \widetilde{\varepsilon}_{ri}^{1/m}\phi_i\right)^m \qquad [75]$$

where $-1 \le 1/m \le 1$ and each phase has a relative permittivity $\widetilde{\varepsilon}_{ri}$ and a volume fraction ϕ_i. This equation simulates mixing of permittivities between the two extremes where the mixed phases are arranged in series ($1/m = -1$) to parallel ($1/m = 1$), with the CRIM model being represented by the value $1/m = 0.5$, that is, $m = 2$. When using this equation, it should be realized that the value of m may change with frequency. When used at high frequencies, in dielectric logging, a polarization factor p for the water phase is often included. In this case, the exponent $1/m$ becomes p/m, but only for the term representing the water phase; the exponents of the terms for the solid and hydrocarbon phases remain $1/m$.

11.04.5.6 Modeling the Frequency-Dependent Electrical Properties of Rocks

There is currently no theoretical model that can predict the frequency-dependent electrical properties of rocks, which contrasts with the models for steady-state conductivity (Archie, 1942; Waxman and Smits, 1968), surface conductivity (Revil and Glover, 1997), zeta potential (Revil et al., 1998), and streaming potential coefficients (Glover et al., 2012a).

The traditional approach has been to match experimental measurements to equivalent electrical circuits composed of capacitors for charge storage and resistors for energy dissipating electrical conduction (e.g., Guéguen and Palciauskas, 1994; Knight et al., 1985; Barsoukov and Macdonald, 2005). Many such equivalent circuits exist, but it is unclear whether any of them help further the understanding of the real physical process occurring in the rock.

Recent theoretical studies have concentrated on characterizing the electrical behavior of geomaterials as macroscopic multicomponent systems over a wide frequency range by combining different influencing polarization mechanisms (e.g., de Lima and Sharma, 1992; Glover et al., 2006b; Leroy and Revil, 2009; Leroy et al., 2008). Often, these models have been compared to experiments that have measured the frequency-dependent electrical properties of glass bead packs. Glover et al. (2006b) combined Debye and Cole and Cole responses for the permittivity of solid grains, pore fluid, and the fixed and diffuse layers of an electrical double layer with a conformable mapping of the Bruggeman–Hanai–Sen effective medium equation (Glover et al., 2010) to predict the frequency response of a pack of glass beads of single size with reasonable results. A much more convincing study was carried out by Leroy et al. (2008) and Leroy and Revil (2009) who used a triple-layer model, and were not only able to model the Maxwell–Wagner response of a bead pack with multiple bead diameters, but were

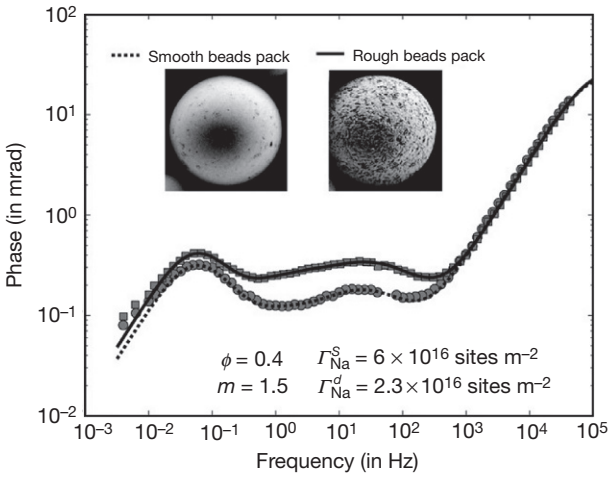

Smooth beads pack
Rough beads pack
Phase (in mrad)
Frequency (in Hz)
$\phi = 0.4$ $\Gamma^S_{Na} = 6 \times 10^{16}$ sites m^{-2}
$m = 1.5$ $\Gamma^d_{Na} = 2.3 \times 10^{16}$ sites m^{-2}

Figure 19 Phase for smooth and rough glass beads (samples 3 and 4) as a function of frequency (from Leroy et al., 2008). The phase is defined here as $\varphi = \tan^{-1}(\sigma''/\sigma')$. The measured porosity is 0.40 and the measured pH is 9.3 for an electrical conductivity of 2.6×10^{-2} S m^{-1}, assuming that $m = 1.5$, with other parameters determined from a TLM model. The modal grain size diameter was 187.5 μm. Symbols: experiment. Lines: model. The left-hand peak is defined by polarization of the triple layer and the size of the beads, the middle peak, by the surface roughness and the right-hand increase by the Maxwell–Wagner effect.

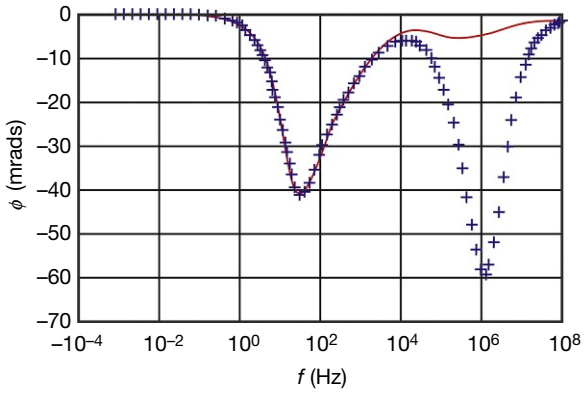


Figure 20 Phase spectrum (dependence of the resistivity phase ϕ_ρ on the frequency f) of a pore model with electrolyte concentration $C_f = 1$ mol m^{-3}, radius of large pore $R_{large} = 1.4 \times 10^{-7}$ m, radius of small pore R_{small} 1.4×10^{-8} m, length of large pore $L_{large} = 10$ μm, and length of small pore $L_{small} = 0.01$ μm; model results without space charges (red solid line) and with space charges allowed (blue crosses). Modified from Volkmann J and Klitzsch N (2010) Frequency-dependent electric properties of microscale rock models for frequencies from one millihertz to ten kilohertz. *Vadose Zone Journal* 9: 858–870, http://dx.doi.org/10.2136/vzj2009.0162.

also able to take account of the polarization due to roughness on the bead surfaces (**Figure 19**).

By contrast, Volkmann and Klitzsch (2010) had focused on the microscopic effects that cause the polarization at the scale of single grains or pores. In 1962, Schwarz (1962) studied the dielectric dispersion of spherical particles in an electrolyte solution and found a Debye-type electrical dispersion with a relaxation time τ that depends on the radius of the particle R:

$$\tau = \frac{R^2}{2\mu_m kT} \qquad [76]$$

where μ_m is the mechanical surface mobility (in s kg^{-1}) of the counterions. If one makes an analogy with rocks, one might say that dispersion processes that occur at a small scale are associated with small relaxation times (i.e., high frequencies) because the spatial influence of a process depends upon the duration of the perturbation for any given mobility and vice versa.

In 1959, Marshall and Madden (1959) developed 'membrane polarization' as a mechanism for describing the mechanism of space charge polarization in geomaterials. Their one-dimensional model consisted of a pore system of alternating 'active' and 'passive' zones. The 'active' zones represent pore throat passages and are characterized by a reduced anion transport number. As these active zones block anions, they behave as ion-selective membranes. For small excess electrolyte concentrations and electric fields, Marshall and Madden (1959) solved the dependence of the complex impedance as a function of the angular frequency and discovered that the relaxation time is related approximately to the square of the length of the active and passive pores. An improvement in the model by Titov et al. (2002) reduced the width and length of the active zones compared to the passive pores and found that the resulting relaxation time is related to the square of the length of the short narrow active zones. Hence, all three theories suggest a quadratic dependence of relaxation time on some geometric length scale of the porous medium, such as pore throat size, grain size, or pore size. This is analogous to the case for fluid permeability (Walker and Glover, 2010) and poses the question whether there is a direct relationship between the relaxation time(s) for electrical dispersion in a geomaterial and its fluid permeability. The prediction that relaxation times are quadratically related to some geometric length scale has been confirmed qualitatively by some experimental data. For example, Kruschwitz (2007) found an exponent of 2.1 for the dependence of the Cole–Cole relaxation time τ_{cc} on the dominant pore throat diameter. Unfortunately, there are plenty of other experimental data that indicate the exponent may be as low as 1.04 (Binley et al., 2005, 2010) to about 3 (Kruschwitz et al., 2010) or even take a logarithmic form in place of a power law (Scott and Barker, 2003, 2005; Titov et al., 2010).

Volkmann and Klitzsch (2010) also noted that there is some evidence for a relationship with the conductivity of the pore fluid. In order to improve the modeling of the membrane polarization approach, Volkmann and Klitzsch (2010) implemented it in three different geometric models and included an electrical double layer. Their finite element modeling showed there are two dispersion mechanisms at work: one associated with a phase angle minimum at a high frequency (approximately 10^5–10^7 Hz) and one at lower frequencies (10^{-3}–10^4 Hz) as shown in **Figure 20**. The high-frequency minimum is associated with the charge buildup at the interface between materials with different electrical properties. This process is known as the Maxwell–Wagner polarization and is described by MWBH theory, which is named after Maxwell (1873), Wagner (1914), Bruggeman (1935), and Hanai (1960). Volkmann and Klitzsch (2010) demonstrated that the high-frequency minimum arose from separation of interfacial space charges by imposing electroneutrality in their model, which removed the space charges and which also caused the high-frequency minimum to be almost

completely removed, as shown in Figure 20. Imposition of electroneutrality did not affect the low-frequency minimum. Volkmann and Klitzsch (2010) attributed this minimum to a 'diffusion current effect'. Its relationship, if any, to the electrochemical Stern layer polarizations studied by Leroy et al. (2008) is not currently clear.

The results of the Volkmann and Klitzsch (2010) model are many and varied. However, the main ones are as follows:

- The only significant geometric length scale to affect the frequency of the low-frequency minimum is the size of pores that are interconnected by narrow pore throats in the direction of current flow.
- The frequency is related to that length scale by a power law with an exponent of either -1.7 or -1.8 depending on the model used (i.e., an exponent of +1.7 and +1.8 relating relaxation time to the geometric length scale).
- There is no relationship between the frequency of the minimum and any other geometric property of the model such as pore throat size and length, such that the characteristic frequency of the process is solely controlled by the size of the pores of the geomaterial in the direction of current flow.
- The phase magnitude of the characteristic frequency is inversely related to the concentration of the pore fluid.

The works of both Leroy et al. (2008) and Volkmann and Klitzsch (2010) involve measurement and modeling of electrical dispersion processes at frequencies less than 200 Hz and sometimes significantly lower. It is becoming clear, therefore, that the control or elimination of electrode polarization in experimental measurements is critical if physical processes occurring in geomaterials are not to be overwhelmed and missed.

Meanwhile, there has also been significant progress in understanding electrical dispersion in colloids, which also have an electrical double layer (Grosse, 2010, 2012). A thorough review is given in Grosse and Delgado (2010).

11.04.5.7 Measurement of Frequency-Dependent Electrical Properties

The measurement of complex conductance, admittance, or permittivity of a geomaterial is usually carried out by impedance spectroscopy. Good treatments of this measurement technique are available (e.g., Lasia, 1999; Barsoukov and Macdonald, 2005; Orazem and Tribollet, 2011; Kanoun, 2012; Macdonald and Kenan, 1987). Measurements are carried out with commercial impedance spectrometers that range from the relatively cheap to the extremely expensive depending on the required specifications, the most important of which is usually the frequency range. However, cheaper but high-quality logging boards that sit inside the logging computer are a relatively recent, attractive alternative. The measurements are complicated by a number of technical issues such as electrode polarization, ground loops, and stray capacitances.

The problem of electrode polarization (Figure 16) can be neglected in rocks above about 1 kHz. For frequencies lower than this, it can be removed completely by using four electrodes. In this arrangement, the measurement of the potential difference across the sample is carried out across the inner pair of electrodes. Since these electrodes carry a negligible current, they do not polarize significantly. All polarization occurs at the outer current electrodes, but since it is not necessary to measure the potential between these two electrodes, their polarization is not important. Sometimes, it can be difficult to incorporate four electrodes into a measurement setup, in a triaxial rig, for example. In this case, a two-electrode system using blacked platinum gauze electrodes has been used successfully (Glover et al., 1994). This paper also describes how blacked platinum electrodes are prepared. In this case, the platinum gauze is electrodeposited with a dendritic coating of amorphous platinum that has a surface area much larger than that of the original platinum gauze. The larger area facilitates the charge transfer equations such as [55], which reduces polarization. Another method is to use an electrode that will not polarize such as an Ag/AgCl electrode (Guichet et al., 2003; Tardif et al., 2011). These electrodes are useful for ambient measurements, but are usually too delicate to be incorporated in cells designed to work at elevated stresses or temperatures.

Careful note should be taken of how the impedance spectrometer is connected to the sample in order to ensure that there are no ground loops or stray capacitances and that electrical noise is minimized. The latter is usually carried out by ensuring that all leads are coaxial and that all coaxial leads are matched to the impedance of the measuring device.

11.04.5.8 Dielectric Properties of Pore Fluids

All gases and crude oils have very low relative permittivity values at all frequencies (less than 2.5), while air is 1.0006 (Lide, 2012). Their permittivities are defined solely by the electronic contribution to polarization.

Aqueous pore fluids, however, have a significant permittivity that is a function of frequency, temperature, and salinity. Water is an extremely polar molecule that follows a Debye frequency dependence with a relaxation time of about 10^{-10} s. At frequencies below the critical frequency, the permittivity depends significantly on temperature and salinity. Higher temperatures reduce the permittivity because thermal agitation opposes the alignment of polar molecules, while the onset of freezing also reduces the permittivity considerably again because polar molecules cannot align themselves in the solid ice. The addition of any solute also reduces the permittivity, but the effect is small. Pure water has a relative permittivity of about 80 (78.36 at 0 Hz and 25 °C (Archer and Wang, 1990)). A table of the complex relative dielectric constants for water in the frequency range 0–50 GHz and temperatures from 0 to 50 °C from Archer and Wang (1990) can also be found in the CRC Handbook of Chemistry and Physics (Lide, 2012). Ice has a typical relative permittivity of 3.1.

A useful low-frequency ($<10^8$ Hz) empirical equation for the relative permittivity of water is given by Olhoeft (Revil et al., 1999b):

$$\varepsilon_r(T, C_f) = a_0 + a_1 T + a_2 T^2 + a_3 T^3 + c_1 C_f + c_2 C_f^2 + c_3 C_f^3 \quad [77]$$

where $a_o = 295.68$, $a_1 = -1.2283\ \mathrm{K}^{-1}$, $a_2 = 2.094 \times 10^{-3}\ \mathrm{K}^{-2}$, $a_3 = -1.41 \times 10^{-6}\ \mathrm{K}^{-3}$, $c_1 = -13.00\ \mathrm{dm}^3\ \mathrm{mol}^{-1}$, $c_2 = 1.065\ (\mathrm{dm}^3\ \mathrm{mol}^{-1})^2$, $c_3 = -0.03006\ (\mathrm{dm}^3\ \mathrm{mol}^{-1})^3$, T is in kelvin and the equation is valid in the range from 273 K to 373 K, and C_f is the salinity of the bulk pore fluid in mol l^{-1}.

Malmberg and Maryott (1956) provided another empirical equation, but for pure water. Their relationship can be written

as $\varepsilon_r(T) = 87.740 - 0.40008T + 9.398 \times 10^{-4}T^2 - 1.410 \times 10^{-6}T^3$, where the temperature is in °C.

11.04.5.9 Dielectric Properties of Minerals and Near-Surface Rocks

11.04.5.9.1 Minerals

The relative dielectric permittivities of most common rock-forming minerals occupy the range 4–10. Since these values are small compared to the dielectric constant of water (about 80) at surface temperatures and pressures, and for frequencies below about 1 GHz, there is little difference in dielectric response between rocks with a quartz, calcite, or dolomite matrix. Additionally, accessory minerals do not normally affect the dielectric response of a rock unless the mineral has a particularly high relative dielectric permittivity (e.g., rutile or pyrite) and the mineral is present in a significant volume fraction. Table 5 shows the relative dielectric permittivities of some of the most common minerals.

11.04.5.9.2 Rocks

Dipolar and interfacial or Maxwell–Wagner polarizations have a much bigger effect on the complex permittivity of porous and fractured rocks than the electronic and atomic polarizations that arise from the minerals they contain. Hence, one would expect the complex electrical properties of rocks to be much more sensitive to changes in pore structure and the saturation of that pore structure with water than to changes in matrix mineralogy.

Effect of pore structure. The dielectric response of geomaterials at frequencies below 1 MHz is related to the amount of water-wetted solid surface present (Li et al., 2001). Increasing the surface area to bulk volume ratio of a geomaterial increases the dielectric response, while treatment of water-wet rock surfaces to make them hydrophobic (Knight and Abad, 1995) removes the dielectric response associated with the development of bound-water layers. In fact, the out-of-phase conductivity σ'' is only a function of the interfacial properties of the interconnected pore surface area. There now exists a substantial volume of literature demonstrating a power-law dependence of σ'' on S_{por}, the surface area to pore volume ratio (Börner and Schön, 1991; Börner et al., 1996; Knight and Nur, 1987b; Schön, 2004; Slater and Lesmes, 2002; Slater et al., 2006) as shown in Figure 21.

Effect of water saturation. Measurements of the dielectric constant of whole or powdered dry rocks show little frequency dependence in the range 10^{-3}–10^{6} Hz (Knight and Nur, 1984, 1987a). Lack of a polar fluid ensures that there is no dielectric polarization and no interfacial polarization can be developed. By comparison, water-saturated rocks have a strongly frequency-dependent dielectric constant that decreases with frequency. Knight and Nur (1984) identified two different frequency regions that have different power-law frequency dependencies. In the high-frequency region (above 10 kHz up to their experimental maximum frequency of 13 MHz), the dielectric constant increased very quickly at low saturations, followed by a more gradual linear increase at medium to high saturations that can be explained by the greater water fraction alone. Knight and Nur (1987a,b) ascribed the strong increases at low saturations to the roles of bound and free water, which has been confirmed by more recent work (Jones and Or, 2002). The situation is further complicated by differences between imbibition and drainage behaviors that result in different microscopic fluid distributions. Despite the apparent complexity, Knight and Endres (1990) and Endres and Knight (1992) were successful in modeling the experimental data. Ulrich and Slater (2004) carried out experimental induced polarization measurements on rocks as a function of water saturation and concluded that IP parameters are a function of saturation for fluid distribution and saturation history.

Biological processes. Recent work has shown that a range of biochemical processes such as those associated with the breakdown of ethanol (Personna et al., 2013a,b) and other organic

Table 5 Relative dielectric permittivity of various common minerals (after a compilation by Schön (2004) of data from Dortman (1976), Keller (1989), Olhoeft (1981), and Parkhomenko (1965))

Mineral	ε_r	*Mineral*	ε_r
Quartz	4.19–5	Apatite	7.4–11.7
Forsterite, fayalite	6.8	Calcite	6.35–8.7
Olivines	6.8	Dolomite	6.3–8.2
Sillimanite	11	Aragonite	7.4–8.67
Amphiboles, hornblende	4.9–8	Anhydrite	5.7–6.7
Graphite	>81	Barite	7–12.2
Hematite	25	Gypsum	5–11.5
Magnetite	34–81	Muscovite	6.2–8
Rutile	79–173	Biotite	6.2–9.3
Pyrite	34–81	Chlorite	9
Galena	17–81	Illite	10
Sphalerite	6.9–12.1	Kaolinite	9.1–11.8
Halite	5.7–6.2	Feldspar	
Sylvite	4.4–6.2	Microcline	5.48–5.6
Fluorite	6.2–8.5	Orthoclase	4.5–6.2
Anorthite	6.9–7.24	Oligoclase	6.06
Labradorite	5.8–5.87	Albite	5.3–6.95

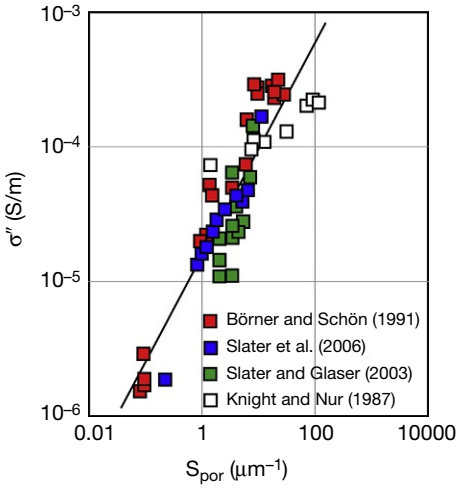


Figure 21 Observed power-law relationship between σ'' and S_{por}. Data from Knight and Nur (1987a,b) and Börner and Schön (1991) are sandstones, data from Slater and Glaser (2003) are alluvial unconsolidated sediments, and data from Slater et al. (2006) are artificial kaolinite–sand mixtures. Figure redrawn after Slater (2007). The linear regression is $\sigma'' = S_{por}^{0.8}$ ($R^2 = 0.82$).

compounds and the biomediated precipitation of metal sulfides (Placencia-Gómez et al., 2013; Zhang et al., 2013) and calcite can produce electrical dispersions with characteristic frequencies often in the low-frequency range (Personna et al., 2008; Zhang et al., 2010, 2012). This work shows the ability of impedance spectroscopy and spectral-induced polarization measurements to monitor the formation of ionic species in groundwater, during bioremediation (Kemna et al., 2012; Mewafy et al., 2013).

Effect of pH. In recent studies, Revil and his colleagues have argued for a dominating role of the Stern layer polarization below 1 kHz in sands and glass beads (Revil and Florsch, 2010; Schmutz et al., 2010; Vaudelet et al., 2011) with counterions like weakly sorbed sodium (sites >Si−O−Na+, where > represents the crystalline framework) being mobile along the mineral surface. Skold et al. (2011) studied the induced polarization as a function of pH both experimentally and by modeling. Their results showed that another polarization mechanism may exist that is associated with a hopping process of the protons along the mineral surface (a so-called Grotthuss cooperation mechanism). Consequently, low-frequency dispersion seems to have many possible contributors associated with the rock fluid interface and the rock microstructure.

Effect of pore fluid salinity and cation type. There is, as expected, a linear relationship between the conductivity of the pore fluid and the real component of electrical conductivity (σ') of the saturated rocks (e.g., Weller et al., 2011). By contrast, the imaginary component (σ'') exhibits a steeper increase at lower salinities that flattens at higher salinities. Weller et al. (2011) also found that sodium cations result in larger values of σ'' than calcium cations in solutions of equal concentration.

In 2011 Revil and Skold (2011) published the results of a combined analytical and experimental study of the salinity dependence of spectral induced polarization in sands and sandstones. They demonstrated analytically that the polarization of the inner part of the electrical triple layer coating the surface of the grains, which is usually called the Stern layer, was consistent with available data. Their model showed the both the specific surface conductivity of the Stern layer and the quadrature conductivity of the porous material depend on the conductivity of the poor water. Once again processes happening very close to the surface of the minerals, that is, in the Stern layer are extremely important to the charge storage properties, and hence frequency-dependent electrical flow in saturated rocks.

Recently, there has been a resurgence in the examination of the effects of the sorption of ions by saturated sands and the consequent effect on induced polarization (Vaudelet et al., 2011a,b). These studies recognize that polarization occurs at frequencies less than a hundred hertz, the most important of which is caused by processes occurring in the electrical double layer, and in the Stern layer in particular. These studies also show that particular cations, zinc and lead for example, can have a large effect on the induced polarization phase angle.

Effect of oil. Since the dielectric constants of gas and oil (about 1 and 2.2, respectively) are so much lower than that of water, dielectric measurements have often been used to monitor the relative saturations of hydrocarbon or water in geomaterials. For example, Vanhala (1997) used induced polarization to map oil-contaminated sand and till. However, the presence of oil in an otherwise water-wet porous rock has been found to have little effect on the frequency dependence of the dielectric response of clean water-wet sand with a relatively low internal surface area ($0.2\ m^2\ g^{-1}$) but has produced decreases as high as 50% for kaolinite, which had considerably larger internal surface area ($5–12\ m^2\ g^{-1}$) (Li et al., 2001). They concluded from their experimental study that the presence of oil leads to a detectable change in the dielectric properties of high-surface-area geomaterials. Recent studies (Schmutz et al., 2010), using the Revil and Florsch (2010) model and applying it to a partially saturated rock found that it was possible to obtain a clear peak at very low frequencies that was associated with the polarization of the Stern layer.

11.04.6 Electrokinetic Properties

Streaming potentials and electroosmosis are examples of coupled properties that link the passage of fluid and electrical flow in reservoir rocks (Jouniaux and Ishido, 2012). These phenomena occur over a wide range of spatial and temporal scales in all porous media that contain ionic fluids (e.g., Figure 2 of Guichet et al. (2003)).

11.04.6.1 Streaming Potentials

One of the fundamental electrokinetic properties is the streaming potential, which is the electric potential that develops when an aqueous fluid flows through a rock. The effect is characterized by a streaming potential coupling coefficient C_s (which is sometimes called the streaming potential coefficient). The streaming potential coefficient is the streaming potential ΔV generated (in V or mV) per difference in fluid pressure ΔP driving the flow (usually in MPa), that is, $C_s = \Delta V/\Delta P$.

Since it is the flow that separates the charge in the porous medium, it has been argued that a better measure may be the streaming potential per flow rate Q, which we call the streaming potential flow coefficient $C_{sv} = \Delta V/Q$ (and then use streaming potential pressure coefficient C_{sp} in place of streaming potential coupling coefficient). While the distinction is not very useful for steady-state conditions, it becomes important for frequency-dependent measurements where the relationship between flow rate and fluid pressure difference also depends upon frequency (see **Chapter 11.02**) and hence allows frequency-dependent processes due to the dynamic flow and the subsequent electrokinetic coupling to be separated.

The mechanism that generates streaming potential arises in the diffuse layer of the electrical double or triple layer (**Section 11.04.4.1**). When fluid flows in the pores of a geomaterial, there exists an interface between the mobile fluid in the body of the pores and that which remains linked to the surface. The interface is called the shear plane. In general, the shear plane occurs in the diffuse layer.

The electric potential on the shear plane is called the zeta potential ζ (in V), which is a parameter of fundamental importance in the theory of electrokinetics. Since, for most rocks, fluid salinities, and compositions, the diffuse layer contains more cations than anions, movement of the electrically neutral bulk pore fluid and a part of the positively charged diffuse layer will cause a positive charge to build up, and so a positive electrical streaming potential is developed in the direction of flow, that is, in the opposite direction to the fluid pressure gradient. Hence, the streaming potential coefficient is generally negative, providing the net charge on the mineral surface is negative.

The streaming potential induces a secondary current in the opposite direction, which is called the streaming current. We have seen [48] that the thickness of the double layer depends upon fluid salinity. However, salinity does not greatly affect the viscosity of the pore fluid. The result is that lower-salinity fluids have wider diffuse layers, but the position of the shear plane hardly moves, ensuring that the separation of charge is more efficient and streaming potentials are greater.

It is generally considered (Wall, 2010) that the first person to carry out experiments involving electroosmosis and electrophoresis was Reuβ (1809), which was followed in 1859 by the first measurement of a streaming potential by Georg Hermann Quincke (Quincke, 1859). Some 20 years later, in 1879, Helmholtz developed a mathematical expression for the effect (Helmholtz, 1879), which was improved by Smoluchowski in 1903 (Smoluchowski, 1903). In the following years, the phenomenon was studied by a number of researchers, including Clark in 1877, Saxen in 1892, and Bikerman in 1932.

11.04.6.2 Electroosmosis

Electroosmosis is the movement of liquid induced by an applied electric potential across a porous material, capillary tube, membrane, microchannel, or any other fluid conduit. Because electroosmotic velocities are independent of conduit size, as long as the double layer is much smaller than the characteristic length scale of the channel, electroosmotic flow is most significant when in small channels. Although the effect is only of the order of a few millimeters per second, it has been used effectively for remediation of pollutants, dewatering, and chemical separation.

The process is formally the opposite of that that gives rise to streaming potentials. The applied electric potential difference acts upon that part of the charged diffuse layer that is mobile, leading to a gross fluid volume flow. Electroosmotic flow was reported first in 1809 by the Russian imperial scientist Ferdinand Friedrich von Reuβ (Reuβ, 1809), who made water to flow through a plug of clay by applying an electric voltage. Although Reuβ studied electrokinetic properties in detail, much was lost to fire in 1812 during Napoleon's occupation of Moscow.

11.04.6.3 Seismoelectrics

Since fluid can be made to flow through rocks locally by the passage of elastic waves, we may also consider that a seismic wavelet may also create an electrical streaming potential. This effect, sometimes called the seismoelectric effect, has components that arise during the passage of the wavelet through the body of the rock and at the interface between rocks of two different properties. The heart of the seismoelectric effect is the coupling between Biot's equation and Maxwell's equations, which was made by Pride (1994) and which is key to the development of the new electroseismic exploration technique (e.g., Glover and Jackson, 2010). In this technique, a layered Earth is perturbed with a seismic pulse; then the resulting electric signals are measured as a function of offset (e.g., Haines et al., 2007; Thompson and Gist, 1993). This method has recently been used to image the vadose zone of a sand aquifer successfully (Dupuis et al., 2007).

11.04.6.4 The Helmholtz–Smoluchowski Equation in the Steady State

The most common approach to modeling the streaming potential is called the Helmholtz–Smoluchowski (HS) equation (e.g., Overbeek, 1952; Hunter, 1981; Maineult et al., 2004; Saunders et al., 2008), which was developed for bundles of capillary tubes. In its simplest form, it is given by

$$C_{sp} = \frac{\Delta V}{\Delta P} = \frac{\varepsilon_r \varepsilon_o \zeta}{\eta_f \sigma_f^*} \qquad [78]$$

where the streaming potential C_{sp} (in V m^{-1}) is the ratio of the measured streaming potential ΔV (in V) to the applied fluid pressure difference ΔP (in Pa) that drives the fluid through the capillary tube. This value depends upon the electric permittivity of the pore fluid $\varepsilon_f = \varepsilon_r \varepsilon_o$ (in F m^{-1}), where ε_r is the relative permittivity of the pore fluid (unitless), ε_o the electric permittivity of free space ($\approx 8.854 \times 10^{-12}$ F m^{-1}), η_f the dynamic viscosity of the pore fluid (in Pa s), σ_f^* the pore fluid electrical conductivity (in S m^{-1}), and ζ the zeta potential (in V). The zeta potential is the electric potential on the shear plane when a part of the diffuse layer is transported by fluid flow.

Equation [78] is commonly applied to porous media including rocks even though it has never been validated for them since there exists no independent method for measuring the zeta potential for complex porous media. The fluid conductivity in eqn [78] has contributions from the bulk fluid and surface conduction. This arises from the derivation of the HS equation. In the derivation, the current caused by ionic transport by advection (i.e., fluid movement) is balanced by a countercurrent, which is conductive. The countercurrent occurs by all conduction mechanisms, which includes conduction through the bulk pore fluid and surface conduction, and would even include conduction through the rock matrix if it were not electrically insulating. Hence, the pore fluid conductivity term in eqn [78] should include a contribution to account for conduction through the bulk fluid and a contribution to account for the surface conduction.

The surface conduction depends on both the surface conductance and the microstructural properties of the rock (porosity, cementation exponent, grain size, etc.) as discussed in detail in Glover and Déry (2010). Consequently, eqn [78] can be rewritten as (Morgan et al., 1989)

$$C_{sp} = \frac{\Delta V}{\Delta P} = \frac{\varepsilon_r \varepsilon_o \zeta}{\eta_f \left(\sigma_f + \frac{2\Sigma_s}{\Lambda} \right)} \qquad [79]$$

where Σ_s is the specific surface conductance (in S) and Λ is a length scale that is characteristic of the pore microstructure (in m) that was introduced by Johnson et al. (1986). This equation implies that the streaming potential coefficient depends on pore size (and consequently grain size and pore throat size) (see **Section 11.04.6.6**). Examination of this implication has led to a model for calculating the zeta potential and streaming potential coefficient for rocks (see **Section 11.04.6.7**).

11.04.6.5 Calculating Zeta Potentials from Experimental Measurements

Equation [78] can be used to derive zeta potentials from experimentally derived streaming potential coefficient data. The pore fluid permittivity can be either measured experimentally or obtained using the empirical relationship given as eqn [77]. The pore fluid viscosity can be measured experimentally or obtained using the empirical relationship developed by Phillips et al. (1978). The pore fluid conductivity must account for any surface conduction that is present. Consequently, eqn [78] is rearranged for the zeta potential with an additional factor to convert the bulk pore fluid conductivity into that that also takes account of the surface conduction. The relevant equation is

$$\zeta = \frac{C_{sp}\eta_f}{\varepsilon_r\varepsilon_o}\,\frac{\sigma_f F_{HS}}{F(C_f)} \qquad [80]$$

where the formation factor F is defined as the ratio of the electrical conductivity of the fluid to that of the fluid-saturated rock, F_{HS} is the formation factor of the sample at high salinity (where the surface conduction is insignificant compared to conduction through the bulk fluid), $F(C_f)$ is the formation factor of the sample at each pore fluid concentration C_f (where there may be a significant contribution of surface conduction to the total fluid conductivity), η_f is the viscosity of the pore fluid, ε_f is the permittivity of the pore fluid, σ_f is the electrical conductivity of the pore fluid, ζ is the zeta potential, and C_{sp} is the measured streaming potential coefficient. The ratio of the two formation factors is termed the Overbeek correction (Overbeek, 1952).

It has also been noted that many streaming potential coefficient data have a quasilinear relationship if the pH is constant and the salinity is not greater than about 1 mol l^{-1} (Jouniaux and Ishido, 2012; Allègre et al. 2010, 2011). A linear fitting gives $C_{sp} = -1.2 \times 10^{-8}/\sigma_f$, which implies a constant zeta potential of about -17 mV (Jouniaux and Ishido, 2012). Recent high-quality measurements where the pH of the pore fluids has been well controlled and for pore fluid salinities from 10^{-5} mol l^{-1} to 6 mol l^{-1} have, however, shown the zeta potential not to be constant, as discussed later.

11.04.6.6 Grain, Pore, and Pore Throat Sizes and Streaming Potential

Equation [79] implies that the streaming potential coefficient is not independent of pore size but diminishes as the characteristic pore scale diminishes. Characteristic pore scale, pore diameter, grain diameter and pore throat diameter are all related. Pore diameter and grain diameter are related by the theta transformation (Glover and Walker, 2009), and the pore size is related to the pore throat size using empirical methods (Glover and Déry, 2010). This last reference uses these transformations to restate [79] in terms of pore diameter, grain diameter, and pore throat diameter for spherical grains as well as grains with arbitrary shape. For example, one relationship for spherical grains is

$$C_{sp} = \frac{\Delta V}{\Delta P} = \frac{d\varepsilon_f\zeta}{\eta_f(d\sigma_f + 6\Sigma_s(F-1))} \qquad [81]$$

while another, which is applicable to grains of different geometries, is

$$C_{sp} = \frac{\Delta V}{\Delta P} = \frac{d\varepsilon_f\zeta}{\eta_f(d\sigma_f + 4\Sigma_s mF)} \qquad [82]$$

Either of these equations can then be used to calculate the streaming coupling coefficient of a rock providing the grain size d (in m), formation factor F, cementation exponent m, and zeta potential (in V) are known. Figure 22 shows the comparison of this model with measurements that were made on random bead packs of different diameters. The dotted line shows the standard HS law, while the grain size models fit the data very well.

11.04.6.7 Streaming Potential and Zeta Potential Models

Equations [81] and [82] and allied equations for pore size and pore throat size may be used to calculate the streaming potential for rocks providing the specific surface conductance and zeta potential can also be modeled. Fortunately, these have been available since Revil and Glover (1997, 1998) and Revil et al. (1999a). Recently, models for all the input parameters to

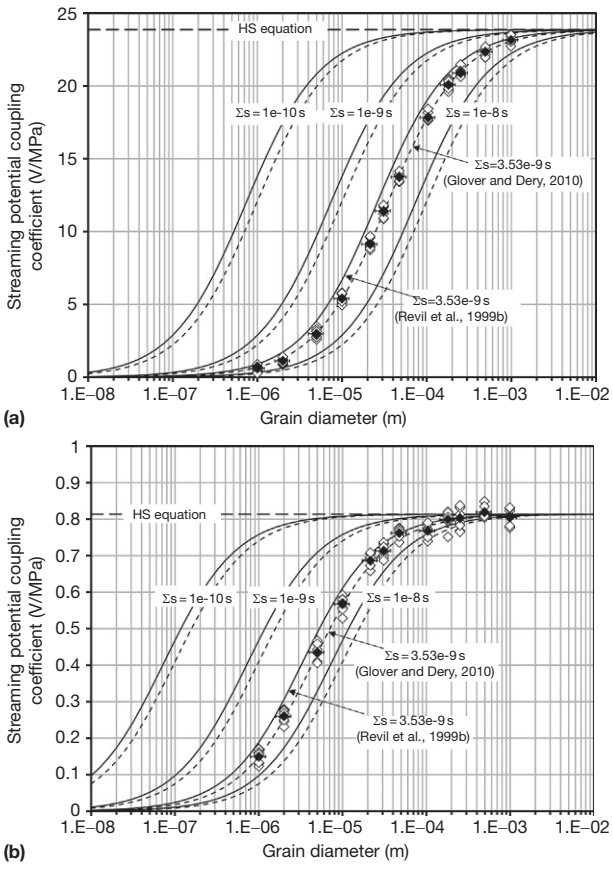


Figure 22 Streaming potential coupling coefficients of a range of quartz glass bead packs as a function of the mean diameter of the glass beads. (a) Pore fluid salinity $C_f = 2 \times 10^{-4}$ mol l^{-1}, (b) pore fluid salinity $C_f = 2 \times 10^{-3}$ mol l^{-1}. The open symbols show the streaming potential coupling coefficient measurements for individual flow rates at a given grain size, which indicates their uncertainty. The solid symbols show the mean value. The horizontal error bar shows the effective standard deviation of the laser diffraction bead size measurements. The solid line shows the model of Revil et al. (1999b) and the dashed line represents that of Glover et al. (2006a,b). Modified from Figure 5 of Glover PWJ and Déry N (2010) Dependence of streaming potential on grain diameter and pore radius for quartz glass beads. *Geophysics* 75: F225–241, http://dx.doi.org/10.1190/1.3509465.

[81] and [82] have been brought together in the first approach that is capable of modeling the streaming potential coefficient of individual rock samples and that takes account of its microstructural properties (Glover et al., 2012a).

Glover et al. (2012a) compared their model for the zeta potential and streaming potential coefficient with 269 zeta potential measurements and 290 streaming potential coefficient measurements for 17 matrix–fluid combinations that were taken from 29 experimental studies in the literature. They found a generally good agreement that became very good if accurate microstructural parameters (grain size, formation factor, and cementation factor) were known. **Figure 23** shows their modeled streaming potentials and zeta potentials when compared against the entire streaming potential coefficient and

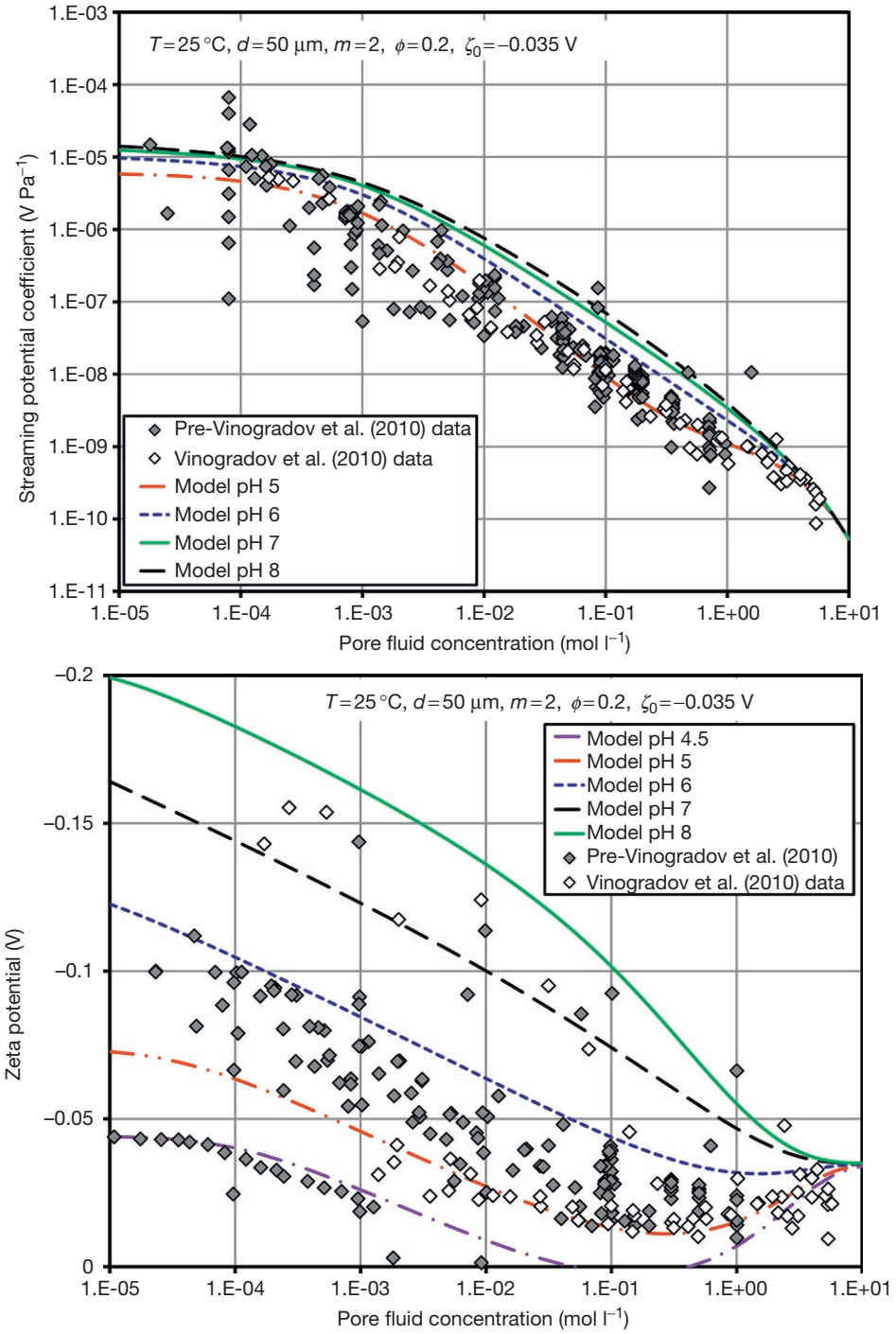


Figure 23 Comparison of the model for streaming potential coefficient (top) and zeta potential (bottom) for a generic rock (grain size = 50 μm, $m = 2$, porosity = 0.2) compared with 290 streaming potential coefficient measurements and 269 zeta potential measurements. Modified from Glover PWJ, Walker E, and Jackson MD (2012) Streaming-potential coefficient of reservoir rock. *Geophysics* 77(2): D17–D43, http://dx.doi.org/10.1190/GEO2011-0364.1.

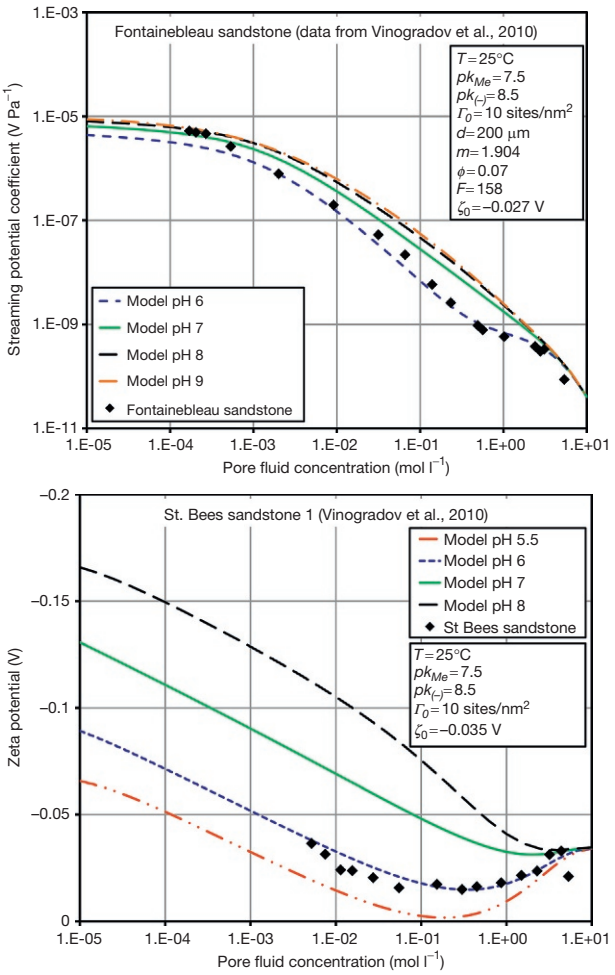


Figure 24 Top: Comparison of the model for the streaming potential coupling coefficient with data measured by Jaafar et al. (2009) for Fontainebleau sandstone (Top) and St. Bees sandstone (Bottom). Modified from Glover PWJ, Walker E, and Jackson MD (2012) Streaming-potential coefficient of reservoir rock. *Geophysics* 77(2): D17–D43, http://dx.doi.org/10.1190/GEO2011-0364.1.

zeta potential experimental datasets, while **Figure 24** shows the theoretical model compared to experimental data from two individual rock samples whose microstructural parameters were well known from independent measurements and were built into the theoretical model so that an individual modeling of each rock sample could be carried out.

It is clear that the model generally follows the trend of all the experimental data for both the streaming potential and the zeta potential, but the fit is much better when individual rock parameters are put into the model.

It is also clear that the pH of the fluid is critical, and since this may change by several points as it becomes equilibrated with the sample, researchers must not only ensure that their measurements take place only after equilibrium of the fluid has been attained but also ensure that the pH (and salinity) is measured on the fluid leaving the sample so that the streaming potential measurements may be attributed to the correct pore fluid salinity and that this salinity is used in the derivation of the zeta potential. The spread in the zeta potential data in **Figure 23** is probably due to variations in the pH of the fluid, and this is being borne out by recent measurements (Walker et al., 2014).

There is much scope for improvement in the model and certain interactions that are not understood. The most important of which is the need to include a zeta potential offset, which is not justified theoretically. However, now that a model exists for calculating the streaming potential coefficient and zeta potentials of rocks, it should be applied routinely to the analysis of all experimental measurements.

11.04.6.8 Factors Affecting the Streaming Potential of Near-Surface Rocks

The model for the steady-state streaming potential coefficient of fully saturated rocks gives us an insight into the many parameters that control electrokinetic conversion in rocks.

The major ones are:

- pore fluid salinity and hence conductivity (e.g., Vinogradov et al., 2010; Walker et al., 2014);
- pore fluid pH (e.g., Ishido and Mizutani, 1981);
- grain, pore, or pore throat size (Glover and Déry, 2010);
- formation factor (e.g., Glover and Déry, 2010; Glover et al., 2012a; Walker et al., 2014);
- cementation exponent (Glover and Déry, 2010; Glover et al., 2012a; Walker et al., 2014);
- porosity (Glover and Déry, 2010; Glover et al., 2012a; Walker et al., 2014);
- the ionic mobility of each of the ions solution (Glover et al., 2012a);
- the disassociation constant of water (e.g., Glover et al., 2012a);
- the equilibrium constant for dissolution of CO_2 in water (Glover et al., 2012a; Revil et al., 1999a; Wu et al., 1991);
- the equilibrium constant for the formation of the carbonate ion in water (Glover et al., 2012a; Revil et al., 1999a; Wu et al., 1991);
- the density of sites on the mineral surface (Glover et al., 2012a; Kosmulski, 1996; Revil and Glover, 1997, 1998);
- the binding constant for cation (sodium) adsorption (Glover et al., 2012a; Kosmulski, 1996; Revil and Glover, 1997; 1998; Scales et al., 1990);
- the disassociation constant for dehydrogenization of silanol (Glover et al., 2012a; Kosmulski, 1996; Revil and Glover, 1997, 1998);
- the distance of the shear plane from the mineral surface (Glover et al., 2012a; Revil and Glover, 1997, 1998);
- the size of protonic surface conduction (Glover et al., 2012a; Revil et al., 1999a);
- the ionic stern plane mobility (Revil and Glover, 1998);
- the dielectric constant of the pore fluid (Glover et al., 2012a; Pride and Morgan, 1991; Walker et al., 2014);
- the viscosity of the pore fluid (Glover et al., 2012a; Pride and Morgan, 1991; Walker et al., 2014); and
- temperature (Reppert and Morgan, 2003a,b), which affects many of the parameters mentioned earlier.

The picture becomes even more complex if one considers a partially saturated rock (**Section 11.04.6.9.7**) or

frequency-dependent streaming potential coefficients (Section 11.04.6.10).

Many of these parameters are either well known for aqueous fluids and common minerals or can be calculated using empirical models. The microstructural parameters and surface conduction are only important for low-salinity fluids and cause the flattening that occurs in the streaming potential coefficient at low salinities (see Figures 23 and 24).

11.04.6.9 Electrokinetic Measurements under Steady-State Conditions

11.04.6.9.1 Experimental considerations

The measurement of streaming potentials, streaming currents, and electroosmotic flows is a large subject that it would be difficult to do justice to in this chapter. The most advanced proponents of these measurements include Laurence Jouniaux (e.g., Allègre et al., 2010; Guichet et al., 2006), Phil Reppert (e.g., Reppert and Morgan, 2001, 2003a,b; Sheffer et al., 2007), André Revil (e.g., Crespy et al., 2007), and Tsuneo Ishido (e.g., Jouniaux and Ishido, 2013). I would recommend that anyone who plans to start these measurements communicate with one of them. It is relatively easy to make a measurement, but it is more challenging to make accurate and well-controlled measurements. We have already discussed how the pore fluid must be in equilibrium with the rock sample so that the relevant salinity and pH are known (Walker et al., 2014), and these must be recorded with the ambient temperature for the measurement.

Analysis of the raw data is important for ensuring their accuracy. Many different ways have been tried, but most often, flow is set up at several different flow rates in one or both directions. The streaming potential coefficient is then calculated from the gradient of the resulting streaming potential/pressure difference graph (Reppert and Morgan, 2001). Flow can be induced by pumping, displacement of fluids by bottled gases, or gravity. Recently, there has been some success using a transient method where a gas bottle displacement has been turned on or a constant flow that is driven by gas displacement has been suddenly stopped (Walker et al., 2014). Logging many thousands of electrical and pressure measurements during the minute or so that the pressure takes to build up or die away allows a very accurate fit to the streaming potential/pressure difference curve to be attained. Measuring the tiny streaming potentials with high-salinity fluids (Vinogradov et al., 2010) or as the pore fluid pH approaches the isoelectric point of the mineral surface (Walker, 2012, pers. comm.) presents the greatest technical difficulties. Since the streaming potentials are steady state, the electrodes have to be nonpolarizing.

Fluid flow is not the only thing that causes electric potentials to develop; they can also arise from redox potentials, be due to temperature differences (thermoelectric effect) (Leinov et al., 2010) or chemical gradients (chemoelectric effect) (Maineult et al., 2005). All of these sources of electric potential should be reduced to a minimum within the experimental rig unless they are the subject of the study.

It is also worth pointing out that microbial activity in the pore fluids can lead to spurious potentials that may confuse electrokinetic measurements. Zhang et al. (2010) made measurements of complex conductivity, electrode potential, and self-potential on a column of sand saturated with river water in which microbial activity was altering the sulfide chemistry of the water. Changes in measured potential due to the production of the bisulfate ion HS^- by sulfate-reducing bacteria reached up to -600 mV compared to a control experiment in which there was no microbial activity. Potential changes started about 6 days after the experiment began and continued until its end, about 22 days later. The changes in potential varied with time and position in the saturated sand column. The measurements show that it is extremely important to exclude biogenic effects from electrokinetic measurements as they can critically confuse the measurements. It may be that the extremely variable potentials measured by Allègre et al. (2010, 2012) may have been affected by this biogeoelectric effect. Conversely, the measurement of this effect may, in future, allow us to monitor the production of certain ionic species in geomaterials using geoelectric measurements. Zhang et al. (2010) inverted their measurements to estimate the biogenic production of bisulfide at each point in their sand column during the whole experiment. This opens up the possibility that the progress of bioremediation of oil spills might be monitored using geoelectric measurements or that microbes might be used as oil-field probes.

For further information on the self-potential technique, please see Chapter 11.09.

11.04.6.9.2 Salinity and pore fluid conductivity

Figure 25 shows typical data for a sandstone from Walker et al. (2014), although here there are more salinities measured than is usually the case, and there are measurements at salinities greater than about 0.3 mol l^{-1}, which are difficult to make. The three samples were all fairly homogeneous and isotropic and show small offsets that are typical of data on rocks from the same lithology and are to small differences in the pH of the equilibrium fluid. The flattening of the data at low salinities is due to the effect of surface conduction on the HS law and is controlled by the specific surface conductance and the microstructural properties of the rock (grain size, formation factor, porosity and cementation exponent). Part of the discrepancy between the datasets arises because each of the three samples equilibrated with the pore fluid at slightly different pH (pH 7, 6, and 6.5, respectively), but all were well modeled using the approach described in Section 11.04.6.7 when the experimental pH had been taken into account. It should be noted that the derived zeta potentials (Figure 25, bottom panel) are certainly not constant with salinity and equal to 17 mV as suggested by Jouniaux and Ishido (2012).

11.04.6.9.3 Pore fluid pH

Measurements as a function of pH are rare because it is difficult to stabilize a solution with a given pH with the rock. Figure 26 shows typical data for silica-based rocks. Here, the data made by Ishido and Mizutani (1981) have been augmented by some very recent data (Walker, 2012, pers. comm.). The pH has a significant effect on the measured streaming potential coefficient and zeta potential, with both approaching to zero near the isoelectric point for silica, $pH_{iso} = 3$ where the electrical double layer collapses because the surface has no net charge.

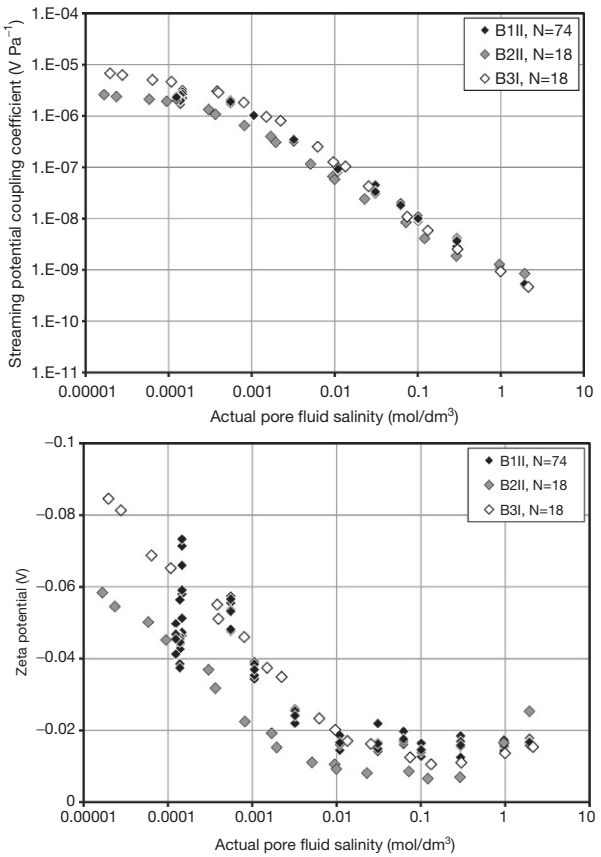


Figure 25 The measured streaming potential coefficient (top) and zeta potential (bottom) for three samples of Boise sandstone as a function of the actual salinity of the aqueous NaCl solution in the pores (B1II, pH = 7.2, T = 26 °C; B2II, pH = 5.94, T = 26 °C; B3I, pH = 6.61, T = 26 °C). Reproduced from Walker E, Glover PWJ, and Ruel J (2014) A transient method for measuring the DC streaming potential coefficient of porous and fractured rocks. *Journal of Geophysical Research: Solid Earth* 119: http://dx.doi.org/10.1002/2013JB010579.

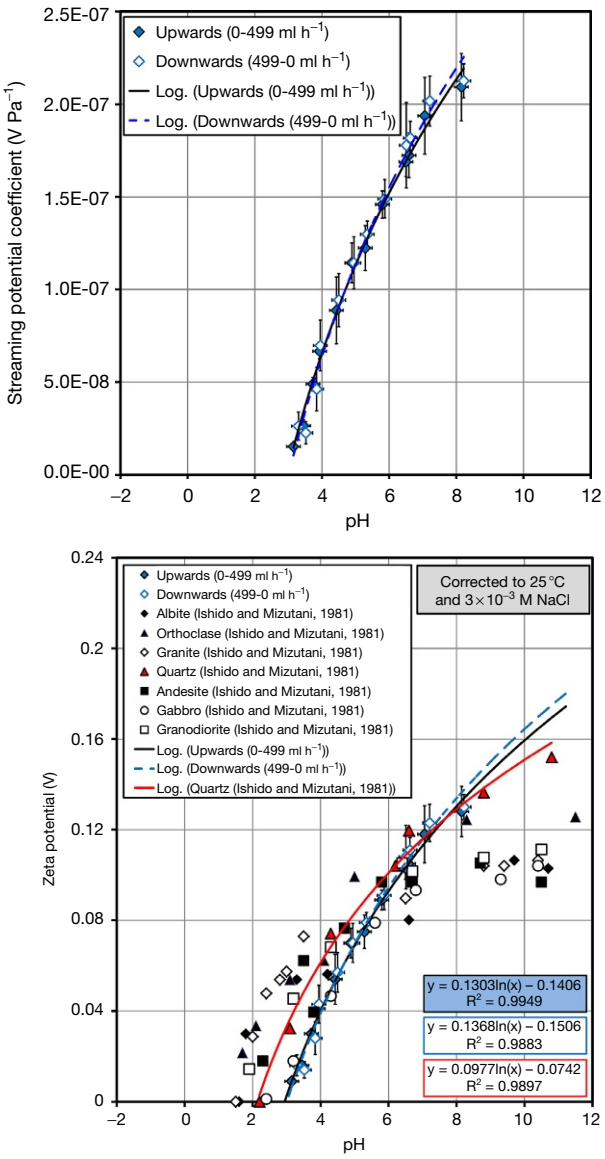


Figure 26 The measured streaming potential coefficient (top) and zeta potential (bottom) as a function of pH for a sample of Berea sandstone saturated with an aqueous NaCl solution (C_f = 3 × 10^{-3} mol l^{-1}, T = 25 °C) compared against data from Ishido and Mizutani (1981) corrected to the same conditions (Walker, 2012, pers. comm.).

11.04.6.9.4 Pore fluid composition

The streaming potential coupling coefficient also depends upon the ionic composition of the pore fluid, and in particular upon the valency of the dominant cation. For silica-based rocks, both Morgan et al. (1989) and Lorne et al. (1999a,b) have shown that fluids whose dominant cation is calcium provide a streaming potential approximately half that of monovalent cations irrespective of salinity. Morgan et al. (1989) tested KOH, KCl, NaCl, HCl, and $CaCl_2$, as a function of salinity, and although there is a clear difference in the streaming potential coefficient decreasing in the order listed, the first four fluids are fairly well grouped, while the divalent cation is significantly less than the others. Lorne tested KCl, K_2SO_4, $CaSO_4$, and $CaCl_2$ in order to examine whether the effect was controlled by divalent cations or anions or both (Lorne et al., 1999a,b). They found that KCl and K_2SO_4 provided the same high values of the coupling coefficient, showing no functional difference between the monovalent Cl^- and the divalent SO_4^{2-}. This is not surprising because for silica-based rocks at their pH (pH = 5.7), the surface is negative and it is the cation that forms the adsorbed layer and is dominant in the diffuse layer. Their divalent cation fluids provided coupling coefficients that were about half as large and once again followed a very similar trend.

11.04.6.9.5 Mineralogy

The great majority of the theoretical and experimental research into the electrokinetic properties of geomaterials has been carried out on silica-dominated rocks. However, some data exist for carbonates (Li et al., 1995; Pengra et al., 1999; Revil and Cerepi, 2004; Sprunt et al., 1994), zeolites, and shales (Revil and Leroy, 2001). Ishido and Mizutani (1981) published zeta potential data for quartz, orthoclase, albite, and granite as a function of pH (Figure 26) and found that all the samples behaved in a similar manner to their Inada granite

(37% quartz, 24% alkali feldspar, 33% plagioclase, and 6% biotite (Oda et al., 2002)), occupying values between the highest values that were given by quartz and the lowest, provided by the albite. As always, the study of shales is complicated by the clays, which have slightly lower zeta potentials than quartz. The clay zeta potential is the result of lower surface site densities but larger specific surface areas. The streaming potential of clay-rich rocks is also fairly low compared to quartz because they have augmented surface conduction that increases the denominator of the HS equation.

11.04.6.9.6 Rock microstructure

There have been no experiments dedicated to examining the effect of rock microstructure upon the streaming potential except grain size, pore size, and pore throat size (Glover and Déry, 2010). The rock microstructure should have no effect upon the zeta potential because the zeta potential is a property of the mineral/fluid interface. However, the theoretical model (**Section 11.04.6.7**) indicates that the grain size, pore size, pore throat size, formation factor, porosity, and cementation exponent should all control the size of the streaming potential coupling coefficient at low salinities, and that variations in these properties between samples are the cause of the divergence in experimental measurements recorded at low salinities which can be seen in **Figure 23**. Small grain, pore, and pore throat sizes all reduce the streaming potential coefficient, as do larger formation factors, and hence, smaller porosities and/or larger cementation exponents. These effects can be understood through [72] and [73], amplifying the effect of surface conduction in the rock. If the surface conductance increases, as in the case of clays, the coupling coefficient is also reduced.

11.04.6.9.7 Saturation

What little experimental evidence exists shows that the streaming potential coefficient is highly sensitive to water saturation, with the size of the coefficient decreasing as the water saturation decreases (Allègre et al., 2010, 2012; Guichet et al., 2003; Revil and Cerepi, 2004; Revil et al., 2007). The coefficient approaches zero at approximately the irreducible water saturation S_{wi}. This observation is important in the application of electrokinetics to oil and gas reservoirs, where the water saturation is generally not 100%. The form of the relative streaming potential coefficient curve as a function of water saturation becomes important in the interpretation of the size of the electrokinetic signal in the reservoir.

When a second fluid phase is present in the pore space, the magnitude of the streaming potential coupling coefficient becomes dependent on a number of factors including the wettability of the rock, the composition of the pore fluids, and the saturation of each phase (Jackson, et al., 2010; Saunders et al., 2012). According to Saunders et al. (2012), in the simplest case, the nonwetting phase is nonpolar and its flux causes no streaming current. The streaming potential coupling coefficient then depends on the saturation of the wetting phase (which we assume to be aqueous) and may be expressed in the form

$$\frac{\Delta V}{\Delta P} = C_i C_{sp} \qquad [83]$$

where the ratio of the streaming potential to the pressure difference that causes fluid flow is equal to the streaming potential coefficient for a fully saturated rock, C_{sp}, multiplied by a relative streaming potential coefficient function, C_i, that accounts for the variation of the streaming potential with water saturation, S_w.

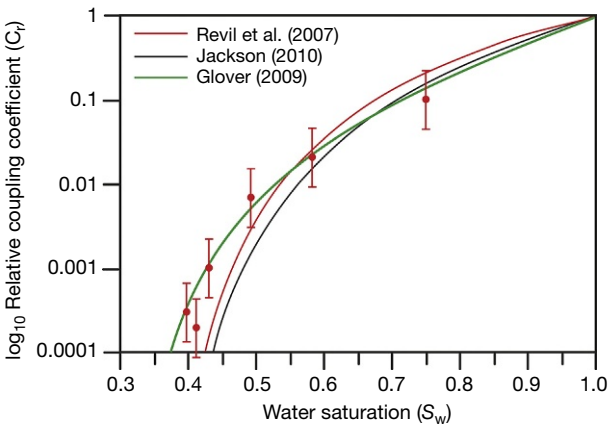


Figure 27 Streaming potential coefficient as a function of water saturation with models (Jackson, 2010; Revil et al., 2007; Glover, 2009, pers. comm.). This diagram uses the values $n = 2.7$, $S_{wi} = 0.36$, and $\lambda = 0.87$ from Revil et al. (2007) and applies them to all the models.

A number of relative streaming potential coefficient functions C_i have been proposed (Revil et al., 2007; Saunders et al., 2012), three of which are given in **Figure 27**.

Revil et al. (2007) suggested that

$$C_i = \frac{1}{S_w^{n+1}} \left(\frac{S_w - S_{wi}}{1 - S_{wi}} \right)^{(2+3\lambda)/\lambda} = \frac{k_{rw}}{S_w^{n+1}} \qquad [84]$$

where k_{rw} is water relative permeability, λ is the Brooks and Corey parameter (Brooks and Corey, 1966), S_w is the water saturation, S_{wi} is the residual water saturation, and n is an exponent.

By contrast, Jackson (2010) used a much simpler approach involving bundles of capillary tubes to obtain

$$C_i = \frac{k_{rw}(S_w)}{\sigma_{rr}(S_w)} = \frac{k_{rw}}{S_w^n} \qquad [85]$$

where the S_w is the volume fraction of water, k_{rw} is the relative permeability of water, and σ_r is the relative electrical conductivity, which is given by Archie's second law as $\sigma_{rr} = S_w^n$ (i.e., the inverse of the resistivity index [29]).

Glover (2009, pers.comm.) had also proposed that

$$C_i = S_w^{n+1} \left(\frac{S_w - S_{wi}}{1 - S_{wi}} \right)^{1/\lambda} \qquad [86]$$

using an empirical approach, which seems to fit the sparse data currently available, but is not based on sound theoretical roots. Other forms of the function exist but much more experimental data are needed in order to discriminate between the models.

The only currently available measurements of the streaming potential coefficient during multiphase flow (Vinogradov and Jackson, 2011) have indicated that there may be a slight hysteresis in the behavior between drainage and imbibition that appears whether the water is replaced by oil (undecane) or gas (moistened nitrogen). In the case of the oil/brine displacement, the streaming potential coefficient relative to that at 100% brine saturation decreased to values between

0.2 and 0.4, but decreased to almost zero for some brine/gas displacements. However, these data are for measurements with saturations at unsteady state and hence need care in their interpretation.

11.04.6.9.8 Temperature

There have been few studies of the dependence of the streaming potential coefficient upon temperature. Those which exist are difficult to interpret because it is often unclear how much of a role is played by thermoelectric effects. Ishido and Mizutani (1981) observed an increase in the streaming potential coefficient with temperature for silica, while Morgan et al. (1989) observed a decrease for Westerly granite.

The variation of streaming potentials and zeta potential with temperature was studied theoretically and experimentally by Reppert and Morgan (2003a,b). They concluded that the variation of streaming potential with temperature cannot be explained solely by the known temperature dependence of water viscosity, permittivity, and conductivity, and that change of zeta potential with temperature must also be taken into account.

Reppert and Morgan (2003a) compiled previous experimental studies, most of which showed that the magnitude of the zeta potential increases with temperature. It was found that this increase is controlled primarily by the surface charge density, which is related to the absorption properties of the surface, the Stern layer charge, and properties of the electrical double layer according to the theory of Revil and Glover (1997).

Reppert and Morgan (2003a) found that the slope of the temperature versus zeta potential curve is controlled by the change in enthalpy of the surface reactions. Overall, the temperature dependence of the zeta potential was found to be determined by the following factors:

- The temperature-dependent behavior of the Debye–Hückel parameter, 4%.
- The temperature-dependent behavior of the diffuse layer of electrical double layer (EDL), 6%.
- The temperature dependence of the surface charge, 90%.

Clearly, it is the surface charge that has the greatest effect.

The streaming potential coefficient also depends upon temperature. The most dominant term controlling the temperature dependence of the streaming potential coefficient is the fluid viscosity.

Reppert and Morgan (2003b) have measured the streaming potential and zeta potential at temperatures between 23 °C and 200 °C and pressures of 20 MPa on samples of Fontainebleau sandstone, Berea sandstone, and Westerly granite. They found that the streaming potential coefficient and zeta potential behaved in a complex fashion that is not completely understood, but probably controlled by multiple mechanisms.

More recently, theoretical modeling (described in **Section 11.04.6.7**) predicts only a small temperature effect that leads to an increase in the magnitude of the streaming potential coefficient and zeta potential up to 120 °C.

11.04.6.10 Streaming Potential as a Function of Frequency

There are surprisingly few measurements of streaming potential for flow that varies in time. This is probably because such measurements have been considered very difficult to carry out. However, such measurements are likely to be of great importance in a number of fields, not least in the understanding and development of the seismoelectric exploration method (Dupuis et al., 2007; Jouniaux and Bordes, 2012).

The few previous measurements can be classified into two groups: (i) transient measurements with a percussive source and (ii) harmonic measurements with a vibrating source (Glover et al., 2012c). The first of these approaches mimics many of the possible applications more closely, while the latter is capable of providing higher-quality frequency-specific data. Glover et al. (2012b) have examined the methods for making frequency-dependent streaming potential measurements, while Reppert and Morgan (2001) have discussed the analysis of the resulting data.

Percussive source studies measure transient electrokinetic processes in sand columns that arise when a controlled impact is made on a column of saturated sand (Bordes et al., 2006, 2008; Chandler, 1981). These are difficult studies that require the impact to be repeatable and the seismoelectric and/or seismomagnetic conversion to be measured. Such transient measurements have confirmed the presence of seismoelectrokinetic and seismoelectromagnetic phenomena at high frequencies. However, the percussive seismic source does not produce a single frequency, so it cannot be used to measure the coupling coefficients as a function of frequency directly. However, it should be possible to examine the Fourier components of the seismic impact and the measured electric and magnetic signals in order to extract more specific information about the frequency dependence of the coupling coefficients (Reppert and Morgan, 2001). Reppert and Morgan carried out this procedure on a sample that they had already measured under steady-state conditions, and by using well-defined single frequencies, producing accurate results from 10 to about 120 Hz in the noninertial flow region. They could not, however, reproduce the decrease in the magnitude of the streaming potential coefficient at frequencies greater than 120 Hz that was measureable using the single-frequency approach.

A better approach is to arrange a harmonically varying fluid flow at a specific frequency and then to measure the time-varying streaming potential and fluid pressure difference, from which a streaming potential coefficient can be calculated. Repeating the experiment at different, well-defined frequencies then allows the variation of the streaming potential coefficient as a function of frequency to be obtained.

There are five ways that such data can be analyzed, and each is described with their advantages and disadvantages by Reppert and Morgan (2001). Briefly, the five approaches are as follows:

- To take the ratio of the RMS (root mean squared) streaming potential to the RMS pressure difference for a train of waves at a given frequency.
- To calculate the ratio of the instantaneous streaming potential to the instantaneous pressure difference for a train of waves at a given frequency and then calculate the mean and standard deviation of the result.
- To plot the instantaneous streaming potential against the instantaneous pressure difference for a train of waves at a given frequency and take the gradient and fitting parameters of the best straight line through the data. This method is

extremely simple but may be complicated if the data are noisy or there is hysteresis (Tardif et al., 2011).

- To cross-correlate both the streaming potential signal and the differential pressure signal, and then to calculate the ratio of the cross-correlated signals. This method works well with noisy data.
- To calculate the amplitude spectrum of both the streaming potential and the differential pressure signals. The streaming potential coefficient is then obtained by taking the ratio of the streaming potential amplitude spectrum to the differential pressure amplitude spectrum at the measured frequency. Other frequencies show the strength of the noise.

There have only been a few AC electrokinetic experiments published in the literature, amounting to measurements on several glass capillaries, one fritted glass filter, and two ceramic filters (Cooke, 1955; Packard, 1953; Reppert, 2000; Reppert and Morgan, 2001; Sears and Groves, 1978). Only one rock has ever been measured at frequencies greater than 100 Hz – a Boise sandstone with 35% porosity (Reppert, 2000). However, recent laboratory developments (Glover et al., 2012b) have provided new apparatuses that have produced high-quality data on Ottawa sand (Tardif et al., 2011) and several grades of glass beads (Glover et al., 2012c) up to 600 Hz, which holds hope for the measurement of rocks in the near future.

Figure 28 shows some recent data for the frequency-dependent streaming potential coefficient made on Ottawa sand together with several models. The top part of this figure shows the magnitude of the complex streaming potential coefficient as a function of frequency, while the bottom two parts show the in-phase (real) and out-of-phase (imaginary) components. All three parts are given because it is important that any model fits the complete complex response. There are a number of models that one can use to fit the data. It is worth noting that none of the models are of the Debye or Cole and Cole type, but all of them are dispersive, with a well-defined critical frequency f_{crit} and a slower roll-off for frequencies greater than the critical frequency, $f > f_{crit}$, than below it, $f < f_{crit}$.

Generic vibrational mechanics models are not normally useful. Sometimes, critically damped or underdamped second-order models are used; but not only do these perform worse than other models, the parameters they provide give no information about the rock. There is some indication that the frequency-dependent behavior of the streaming potential may follow a Cole–Davidson frequency dependence (Davidson and Cole, 1951). Packard (1953) introduced a model based on bundles of capillary tubes that was subsequently simplified (Reppert et al., 2001), noting that Glover et al. (2012c) contains an important correction to a misprint in the Reppert et al. (2001) paper. The best model, however, is

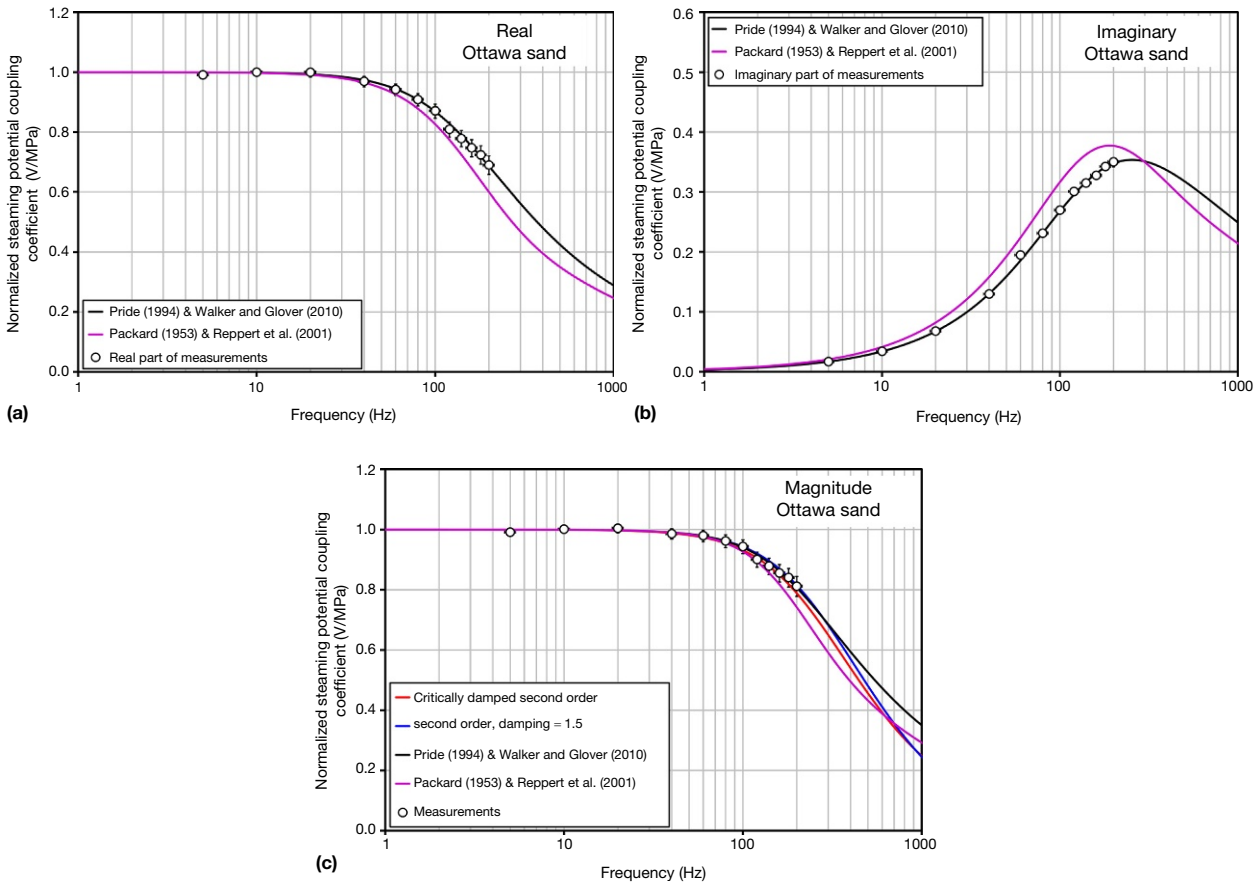


Figure 28 The calculated streaming potential coupling coefficient (normalized to the value at 5 Hz, which was 0.518 V MPa^{-1}) as a function of frequency for Ottawa sand saturated with 10^{-3} mol l^{-1} NaCl solution at 24 °C (porosity = 0.325, $\kappa_{10} = 1.19 \times 10^{-10}$ m^2). (c) Magnitude with six models shown fitting the data, (a) real component with four models, and (b) imaginary component with four models (Packard, 1953; Pride, 1994; Reppert et al., 2001; Thomson and Dahleh, 1998; Walker and Glover, 2010).

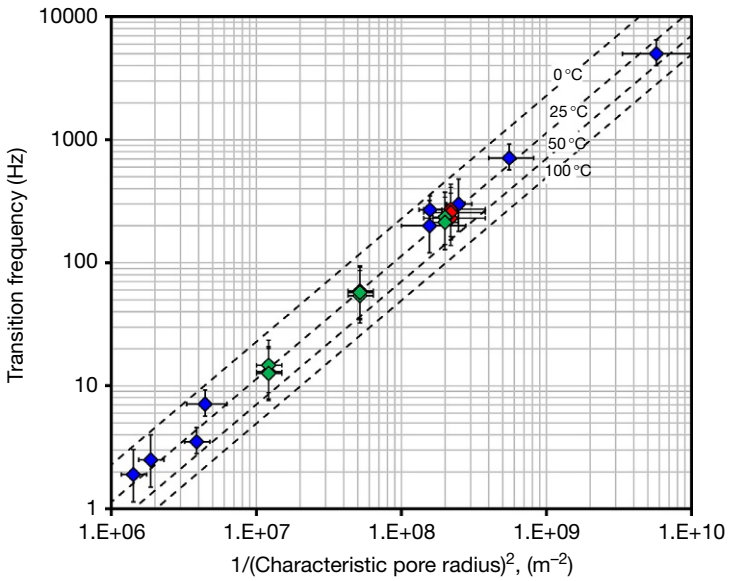


Figure 29 The electrokinetic transition frequency as a function of the inverse square characteristic pore size. Blue symbols, previous data for capillary tubes, filters, frits, and rocks (sources in Walker and Glover, 2010); red symbols, Ottawa sand measured at 24.0 °C ± 0.8 °C (Glover et al., 2012b); green symbols, glass bead packs (Glover et al., 2012b). The dashed lines represent the theoretical result [87] at four different temperatures to take account of the temperature dependence of both the fluid viscosity and the density.

that of Pride (1994), whose model was derived for a porous medium. While this model is complex, Walker and Glover (2010) provided a simplified version.

The critical angular frequency ω_{crit} ($=2\pi f_{crit}$) is related to the inverse square of the effective pore radius of the rock r_{eff} (Pride, 1994; Walker and Glover, 2010):

$$\omega_{crit} = \frac{8}{r_{eff}^2}\frac{\eta_f}{\rho_f} \qquad [87]$$

where η_f and ρ_f are the viscosity and density of the saturating fluid, respectively. Figure 29 shows this relationship for all the data currently in the public domain. Since the effective pore radius of the rock can be related to the permeability of the rock, we may use the critical frequency as a proxy for rock permeability or to predict permeability (Glover et al., 2012a,c).

11.04.7 Summary

11.04.7.1 History and Progress

Twenty years ago the electrical properties of porous and cracked rocks were described using a set of empirical relationships of which those by Archie were preeminent. They were sufficiently successful in the majority of cases that they were used in almost every assessment of oil and gas reserves for the previous 50 years.

In the last two decades, there have been large advances in the theory of the electrical properties of rocks; advances thanks to researchers such as André Revil, Steve Pride, and David Lockner. We now have theoretical tools to predict the electrical behavior of a rock in the DC regime, and at last, progress is being made in the development of theory in the AC regime. We have methods that enable the surface charge on minerals and the resulting surface conduction to be calculated, as well as a theoretical method for calculating the zeta potential. Progress in electrokinetics in both the steady-state and the frequency-dependent regimes allows us to measure and to predict the streaming potential coefficient.

Applications of the electrical and electrokinetic properties of rocks are to be found in many areas of geophysics well outside the traditional application in hydrocarbon exploration. Within the oil and gas industry, the era of large simple reservoirs is long over. Now, reserves are to be found in many, small, heterogeneous, anisotropic, difficult, and unconventional reservoirs. The analysis of these reservoirs will require the new sophisticated theoretical methods as well as the fully trained petrophysicists to use them.

11.04.7.2 Current Challenges and Future Directions

Two competing ideas can be applied to most natural sciences. The first is that any deep scientific truth will be simple and elegant. That may work well in particle physics, but such profound patterns and symmetries are difficult to see in the application of physics to natural materials. It is as if the nice bright shiny theories of Faraday and Maxwell are stained, dirtied, and rather well camouflaged by Mother Nature.

The second idea stems from this realization. It is that the Earth and its constituents are complex, and hence, their physics will be complex. It is worth realizing that the biggest challenge involving geomaterials is the complexity of their microstructure. This microstructure defines the physical properties of the material by acting as incredibly convoluted boundary conditions to the solution of constituent relationships, whether they are electrical, elastic wave, or thermal.

Advances in the speed of computers and the efficacy of programs designed to handle large datasets have started to enable us to untangle the complicated responses of complex

geomaterials and will continue to do so. The untangling of the effects of microstructure may not have the appeal of other research challenges, yet I believe that it is the biggest challenge in petrophysics.

Plenty of challenges remain. Most are related to microstructure. The most pressing challenges are as follows:

(i) To set up a fundamental theoretical framework for the frequency-dependent electrical properties of rocks that does not depend upon standard dispersion mechanisms and that is compatible with a zero frequency solution.
(ii) To ensure that all rock physical properties are viewed in terms of the way that they couple to other properties. Electrical properties couple to hydraulic, elastic wave, chemical, thermal, and magnetic properties, among others.
(iii) To improve the ability to make high-quality measurements, especially the extension of the ranges of frequency for which measurements can usefully be made.
(iv) To combine new experimental techniques with high resolutions, such as combining x-ray computed microtomography (μCT) with advanced modeling techniques to solve locally valid equations for the electrical properties of rocks, and to compare the results with existing empirical, quasiempirical, and effective medium models that are currently used.
(v) To ensure that all theory, measurements, and modeling are immediately and effectively applied to practical challenges with socioeconomic impact.
(vi) In the future, hydrocarbon reservoirs will be increasingly heterogeneous and anisotropic and smaller. The challenge will be to develop simple and effective methods for characterizing and managing production from these smaller more complex reservoirs while increasing the automation of the process. Geoelectric properties will have a part to play.

References

Adamson AW (1976) *Physical Chemistry of Surfaces*. New York: John Wiley and Sons.

Allègre V, Jouniaux L, Lehmann F, and Sailhac P (2010) Streaming potential dependence on water-content in Fontainebleau sand. *Geophysical Journal International* 182(3): 1248–1266. http://dx.doi.org/10.1111/j.1365-246X.2010.04716.x.

Allègre V, Jouniaux L, Lehmann F, and Sailhac P (2011) Reply to comment by A. Revil and N. Linde on 'Streaming potential dependence on water-content in Fontainebleau sand'. *Geophysical Journal International* 186(1): 115–117. http://dx.doi.org/10.1111/j.1365-246X.2011.05038.x.

Allègre V, Lehmann F, Ackerer P, Jouniaux L, and Sailhac P (2012) A 1-D modelling of streaming potential dependence on water content during drainage experiment in sand. *Geophysical Journal International* 189(1): 285–295. http://dx.doi.org/10.1111/j.1365-246X.2012.05371.x.

Archer DG and Wang P (1990) The dielectric constant of water and Debye-Hückel limiting law slopes. *Journal of Physical and Chemical Reference Data* 19(2): 371.

Archie GE (1942) The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of AIME* 146(1): 54–62. http://dx.doi.org/10.2118/942054-G.

Baba K (2005) Electrical structure in marine tectonic settings. *Surveys in Geophysics* 26(6): 701–731.

Banton O, Seguin M, and Cimon M (1997) Mapping field-scale physical properties of soil with electrical resistivity. *Soil Science Society of America Journal* 61(4): 1010–1017.

Bardon C and Pied B (1969) Formation water saturation in shaly sands. In: *Society of Professional Well Log Analysts10th Annual Logging Symposium Transactions, Paper Z*, 19 pp.

Barsoukov E and Macdonald JR (2005) *Impedance Spectroscopy: Theory, Experiment, and Applications*. John Wiley, ISBN: 0471647497, 608 pp.

Beekmans NM and Heyne L (1976) Correlation between impedance, microstructure and compositions of calcia-stabilised zirconia. *Electrochimica Acta* 21: 303–310.

Bentley RW (2002) Global oil and gas depletion: An overview. *Energy Policy* 30(3): 189–205. http://dx.doi.org/10.1016/S0301-4215(01)00144-6.

Bernabé Y and Bruderer C (1998) Effect of the variance of pore size distribution on the transport properties of heterogeneous networks. *Journal of Geophysical Research* 103: 513–526.

Binley A, Slater LD, Fukes M, and Cassiani G (2005) Relationship between spectral induced polarization and hydraulic properties of saturated and unsaturated sandstone. *Water Resources Research* 41: W12417. http://dx.doi.org/10.1029/2005WR004202.

Binley A, Kruschwitz S, Lesmes D, and Kettridge N (2010) Exploiting the temperature effects on low frequency electrical spectra of sandstone: A comparison of effective diffusion path lengths. *Geophysics* 75(6): A43–A46.

Bordes C, Jouniaux L, Dietrich M Pozzi, Pozzi J-P, and Garambois S (2006) First laboratory measurements of seismo-magnetic conversions in fluid-filled Fontainebleau sand. *Geophysical Research Letters* 33: L01302. http://dx.doi.org/10.1029/2005GL024582.

Bordes C, Jouniaux L, Garambois S, Dietrich M Pozzi, Pozzi J-P, and Gaffet S (2008) Evidence of the theoretically predicted seismo-magnetic conversion. *Geophysical Journal International* 174: 489–504. http://dx.doi.org/10.1111/j.1365-246X.2008.03828.x.

Börner FD and Schön JH (1991) A relation between the quadrature component of electrical conductivity and the specific surface area of sedimentary rocks. *The Log Analyst* 32: 612–613.

Börner FD, Schopper JR, and Weller A (1996) Evaluation of transport and storage properties in the soil and groundwater zone from induced polarization measurements. *Geophysical Prospecting* 44: 583–602.

Brooks RH and Corey AT (1966) Properties of porous media affecting fluid flow. *Journal of the Irrigation and Drainage Division of the American Society of Civil Engineering* 92(IR2): 61–88.

Bruggeman DAG (1935) Berechnung verschiedener physikalischer Konstanten von heterogenen Substanzen. I. Dielektrizitätskonstanten und Leitigähigkeiten der Mischkörper aus isotropen Substanzen. *Annalen der Physik* 5: 636–664.

Bussian AE (1983) Electrical conductance in a porous medium. *Geophysics* 1983(48): 1258–1268. http://dx.doi.org/10.1190/1.1441549.

Carter LM, Campbell BA, Holt JW, et al. (2009) Dielectric properties of lava flows west of Ascraeus Mons, Mars. *Geophysical Research Letters* 36(23).

Chandler R (1981) Transient streaming potential measurements on fluid-saturated porous structures: An experimental verification of Biot's slow wave in the quasi-static limit. *Journal of the Acoustical Society of America* 70: 116–121.

Chelidze TL and Guéguen Y (1999) Electrical spectroscopy of porous rocks: A review—I. Theoretical models. *Geophysical Journal International* 137(1): 1–15. http://dx.doi.org/10.1046/j.1365-246X.1999.00799.x.

Chelidze TL, Guéguen Y, and Ruffet C (1999) Electrical spectroscopy of porous rocks: A review—II. Experimental results and interpretation. *Geophysical Journal International* 137(1): 16–34. http://dx.doi.org/10.1046/j.1365-246X.1999.00800.x.

Chelidze T, Varamashvili N, Devidze M, Tchelidze Z, Chikhladze V, and Matcharashvili T (2003) Laboratory study of electromagnetic initiation of slip. *Annali di Geofisica* 45(5): 587–598.

Clavier C, Coates G, and Dumanoir L (1977) The theoretical and experimental bases for the Dual Water Model for the interpretation of shaly sands. In: *Paper SPE 6859 presented at the 1977 SPE Annual Technical Conference and Exhibition, Denver, October 9–12*. http://dx.doi.org/10.2118/6859-PA.

Clennell MB (1997) *Tortuosity: A Guide Through the Maze*. Vol. 122. London: Geological Society, Special Publications, January, pp. 299–344. http://dx.doi.org/10.1144/GSL.SP.1997.122.01.18.

Coates GR, Boutemy Y, and Clavier C (1983) A study of the dual-water model based on log data. *Journal of Petroleum Technology* 35(1): 158–166. http://dx.doi.org/10.2118/10104-PA.

Cole KS and Cole RH (1941) Dispersion and absorption in dielectrics I: Alternating current characteristics. *Journal of Chemical Physics* 9: 341–351. http://dx.doi.org/10.1063/1.1750906.

Cooke CE (1955) Study of electrokinetic effects using sinusoidal pressure and voltage. *Journal of Chemical Physics* 23: 2299–2303. http://dx.doi.org/10.1063/1.1740742.

Crespy A, Bolève A, and Revil A (2007) Influence of the Dukhin and Reynolds numbers on the apparent zeta potential of granular porous media. *Journal of Colloid and Interface Science* 305(1): 188–194. http://dx.doi.org/10.1016/j.jcis.2006.09.038.

Cyr G, Glover PWJ, and Novikov V (2010) Can an electro-kinetic mechanism explain artificial earthquakes? In: *EGU2010, EGU2010-2526, Vienna, 2–7 May 2010*.

Dai L and Karato S (2009) Electrical conductivity of pyrope-rich garnet at high temperature and pressure. *Physics of the Earth and Planetary Interiors* 176: 83–88.

David C (1993) Geometry of flow paths for fluid transport in rocks. *Journal of Geophysical Research* 98: 12267–12278.

Davidson DW and Cole RH (1951) Dielectric relaxation in glycerol, propylene glycol, and n-propanol. *Journal of Chemical Physics* 29: 1484–1490.

Davis J and Kent D (1990) Surface complexation modeling in aqueous geochemistry. In: Hochella MF and White AF (eds.) *Mineral Water Interface Geochemistry*. Mineralogical Society of America.

Davis JA, James RO, and Leckie JO (1978) Surface ionization and complexation at the oxide/water interface. I. Computation of electrical double layer properties in simple electrolytes. *Journal of Colloid and Interface Science* 63(3): 480–499. http://dx.doi.org/10.1016/S0021-9797(78)80009-5.

de Lima OAL and Sharma MM (1992) A generalized Maxwell–Wagner theory for membrane polarization in shaly sands. *Geophysics* 57: 431–440.

Debye PJW (1929) *Polar Molecules*. New York: Initially Chemical Catalogue Co., subsequently Dover Press, pp. 77–108, Chapter 5.

Di Maio R and Patella D (1991) Basic theory of electrokinetic effects associated with earthquakes. *Bollettino di Geofisica Teorica ed Applicata* 33(130–131): 145–154.

Di Maio R and Patella D (1994) Self-potential anomaly generation in volcanic areas. The Mt. Etna case-history. *Acta Vulcanologica* 4: 119–124.

Dissado LA and Hill RM (1984) Anomalous low-frequency dispersion. *Journal of the Chemical Society, Faraday Transactions* 2: 291–319. http://dx.doi.org/10.1039/F29848000291.

Dortman NB (1976) *Fiziceskie svoistva gornich porod I polesnich iskopamych*. Nedra, Moscow: Izdat.

Doveton JH (2001) All models are wrong, but some models are useful: "Solving" the Simandoux equation. In: *Annual Conference of the International Association of Mathematical Geology, Section J, Cancun, Mexico*. http://www.kgs.ku.edu/Conferences/IAMG/Sessions/J/Papers/doveton.pdf.

Duba A, Heikamp S, Meurer W, Mover G, and Will G (1994) Evidence from borehole samples for the role of accessory minerals in lower-crustal conductivity. *Nature* 367: 59–61. http://dx.doi.org/10.1038/367059a0.

Dukhin SS and Derjaguin BV (1974) In: Matijevic E (ed.) *Surface and Colloid Science*. New York: John Wiley and Sons.

Dunlap HF and Hawthorne RR (1951) The calculation of water resistivities from chemical analyses. *Journal of Petroleum Technology* 3(3): 373–375.

Dupuis JC, Butler KE, and Kepic AW (2007) Seismoelectric imaging of the vadose zone of a sand aquifer. *Geophysics* 72(6): A81–A85. http://dx.doi.org/10.1190/1.2773780.

Einaudi F, Pezard PA, Ildefonse B, and Glover P (2005) *Electrical Properties of Slow-Spreading Ridge Gabbros from ODP Hole 1105A, SW Indian Ridge*, vol. 240, pp. 179–193, Geological Society Special Publication. http://dx.doi.org/10.1144/GSL.SP.2005.240.01.14.

Ellis DV and Singer JM (2007) *Well Logging for Earth Scientists*, 2nd ed. Dordrecht, The Netherlands: Springer, ISBN: 9781402037382.

Endres AL and Knight R (1992) A theoretical treatment of the effect of microscopic fluid distribution on the dielectric properties of partially saturated rocks. *Geophysical Prospecting* 40(3): 307–324.

Ewing RP and Hunt AG (2006) Dependence of the electrical conductivity on saturation in real. Porous media. *Vadose Zone Journal* 5(2): 731–741. http://dx.doi.org/10.2136/vzj2005.0107.

Focke JW and Munn D (1987) Cementation exponents in middle eastern carbonate reservoirs. *SPE Formation Evaluation* 2(2): 155–167.

Friedrichs B, Matzander U, Large D, and Davies M (1999) Controlled source electro-magnetic mapping-development and evaluation of its application to mineral exploration. *Transactions of the Institutions of Mining and Metallurgy, Section B: Applied Earth Science* 108(Sept–Dec): B178–B182.

Fujinawa Y, Kumagai T, and Takahashi K (1992) A study of anomalous underground electric field variations associated with a volcanic eruption. *Geophysical Research Letters* 19(1): 9–12.

Gaffney C (2008) Detecting trends in the prediction of the buried past: A review of geophysical techniques in archaeology. *Archaeometry* 50(2): 313–336.

Glover PWJ (1996) Graphite and electrical conductivity in the lower continental crust: A review. *Physics and Chemistry of the Earth* 21(4): 279–287.

Glover PWJ (2009) What is the cementation exponent? A new interpretation. *The Leading Edge:* 82–85. http://dx.doi.org/10.1190/1.3064150.

Glover PWJ (2010a) The discovery of an Anglo-Saxon grubenhaus at New Bewick, Northern UK using electrical surveying and predictive deconvolution. *Archaeometry* 52(2): 320–342.

Glover PWJ (2010b) A generalized Archie's law for *n* phases. *Geophysics* 75(6): E247–E265. http://dx.doi.org/10.1190/1.3509781.

Glover PWJ and Adám A (2008) Correlation between crustal high conductivity zones and seismic activity and the role of carbon during shear deformation. *Journal of Geophysical Research* 113: B12210. http://dx.doi.org/10.1029/2008JB005804.

Glover PWJ and Déry N (2010) Dependence of streaming potential on grain diameter and pore radius for quartz glass beads. *Geophysics* 75: F225–F241. http://dx.doi.org/10.1190/1.3509465.

Glover PWJ and Jackson MD (2010) Borehole electro-kinetics. *The Leading Edge* 29: 724–728. http://dx.doi.org/10.1190/1.3447786.

Glover PWJ and Vine FJ (1992) Electrical conductivity of carbon-bearing granulite at raised temperatures and pressures. *Nature* 360(6406): 723–726.

Glover PWJ and Vine FJ (1994) Electrical conductivity of the continental crust. *Geophysical Research Letters* 21(22): 2357–2360.

Glover PWJ and Vine FJ (1995) Beyond KTB – electrical conductivity of the deep continental crust. *Surveys in Geophysics* 16(1): 5–36.

Glover PWJ, Meredith PG, Sammonds PR, and Murrell SAF (1994) Ionic surface electrical conductivity in sandstone. *Journal of Geophysical Research* 99: 21635–21650. http://dx.doi.org/10.1029/94JB01474.

Glover PWJ, Gomez JB, Meredith PG, Hayashi K, Sammonds PR, and Murrell SAF (1997) Damage of saturated rocks undergoing triaxial deformation using complex electrical conductivity measurements: Experimental results. *Physics and Chemistry of the Earth* 22: 57–61. http://dx.doi.org/10.1016/S0079-1946(97)00078-5.

Glover PWJ, Hole MJ, and Pous J (2000a) A modified Archie's Law for two conducting phases. *Earth and Planetary Science Letters* 180: 369–383. http://dx.doi.org/10.1016/S0012-821X(00)00168-0.

Glover PWJ, Pous J, Queralt P, Muñoz J-A, Liesa M, and Hole MJ (2000b) Integrated two dimensional lithospheric conductivity modelling in the Pyrenees using field-scale and laboratory measurements. *Earth and Planetary Science Letters* 178(1–2): 59–72.

Glover PWJ, Gómez JB, and Meredith PG (2000c) Fracturing in saturated rocks undergoing triaxial deformation using complex electrical conductivity measurements: Experimental study. *Earth and Planetary Science Letters* 183: 201–213. http://dx.doi.org/10.1016/S0012-821X(00)00267-3.

Glover PWJ, Zadjali I, and Frew K (2006a) Permeability prediction from MICP and NMR data using an electro-kinetic approach. *Geophysics* 71: F49.

Glover PWJ, Déry N, and Ransford TJ (2006b) The complex electrical properties of rocks and bead-packs as a function of frequency: Experiment and modelling. In: *Ann. Gen. Ass. EGU 2006, EGU06-A-10732, Vienna*.

Glover PWJ, Ransford TJ, and Auger G (2010) A simple method for solving the Bussian equation for electrical conduction in rocks. *Solid Earth* 1: 85–91. http://dx.doi.org/10.5194/se-1-85-2010.

Glover PWJ and Walker E (2009) Grain-size to effective pore-size transformation derived from electrokinetic theory. *Geophysics* 74(1): E17–E29. http://dx.doi.org/10.1190/1.3033217.

Glover PWJ, Walker E, and Jackson MD (2012a) Streaming-potential coefficient of reservoir rock. *Geophysics* 77(2): D17–D43. http://dx.doi.org/10.1190/GEO2011-0364.1.

Glover PWJ, Ruel J, Tardif E, and Walker E (2012b) Frequency-dependent streaming potential of porous media—Part 1: Experimental approaches and apparatus design. *International Journal of Geophysics* 2012: 846204. http://dx.doi.org/10.1155/2012/846204.

Glover PWJ, Walker E, Ruel J, and Tardif E (2012c) Frequency-dependent streaming potential of porous media—Part 2: Experimental measurement of unconsolidated materials. *International Journal of Geophysics* 2012: 728495. http://dx.doi.org/10.1155/2012/728495.

Green AD (1997) Measurements of the dielectric properties of cheddar cheese. *Journal of Microwave Power and Electromagnetic Energy* 32(1): 16–27.

Grosse C (2010) Extension of a classic theory of the low frequency dielectric dispersion of colloidal suspensions to the high frequency domain. *Journal of Physical Chemistry B* 114(39): 12520–12527.

Grosse C (2012) Extension of a classic low frequency dielectric dispersion theory of colloidal suspensions to include different counterion and co-ion valences, a broad frequency range, and the stagnant layer conductivity. *Journal of Physical Chemistry B* 116(45): 13538–13553.

Grosse C and Delgado AV (2010) Dielectric dispersion in aqueous colloidal systems. *Current Opinion in Colloid and Interface Science* 15(3): 145–159.

Guéguen Y and Palciauskas V (1994) *Introduction to the Physics of Rocks*. Princeton, NJ: Princeton University Press, ISBN: 9780691034522.

Guichet X, Jouniaux L, and Pozzi J-P (2003) Streaming potential of a sand column in partial saturation conditions. *Journal of Geophysical Research, B: Solid Earth* 108(3): ECV 2-1–2-12. http://dx.doi.org/10.1029/2001JB001517.

Guichet X, Jouniaux L, and Catel N (2006) Modification of streaming potential by precipitation of calcite in a sand-water system: Laboratory measurements in the pH range from 4 to 12. *Geophysical Journal International* 166(1): 445–460. http://dx.doi.org/10.1111/j.1365-246X.2006.02922.x.

Gung Y, Panning M, and Romanowicz B (2003) Global anisotropy and the thickness of continents. *Nature* 422(6933): 707–711.

Haines SS, Pride SR, Klemperer SL, and Biondi B (2007) Seismoelectric imaging of shallow targets. *Geophysics* 72(2): G9–G20. http://dx.doi.org/10.1190/1.2428267.

Hanai T (1960) Theory of the dielectric dispersion due to the interfacial polarization and its application to emulsions. *Kolloid-Z*. 171: 23–31.

Hashin Z and Shtrikman S (1962) A variational approach to the theory of effective magnetic permeability of multiphase materials. *Journal of Applied Physics* 33: 3125–3131.

Helmholtz H (1879) Studien über electrische Grenzschichten. *Annalen der Physik* 7: 337–382.

Hilchie DW (1984) A new water resistivity versus temperature equation. *The Log Analyst* 25(4): 20–21.

Hossin A (1960) Calcul des saturations en eau par la methode du ciment argileux (formule d'Archie generalisee). *Bulletin Association Francaise des Techniciens du Petrole* 140.

Hunt AG (2004) Continuum percolation theory and Archie's law. *Geophysical Research Letters* 31(L19503): 1–4.

Hunter RJ (1981) *Zeta Potential in Colloid Science*. New York: Academic Press.

Iler RK (1979) *The Chemistry of Silica*. New York: Wiley.

Ishido T and Mizutani H (1981) Experimental and theoretical basis of electro-kinetic phenomena in rock-water systems and its applications to geophysics. *Journal of Geophysical Research* 86: 1763–1775.

Izgorodina EI, Forsyth M, and MacFarlane DR (2009) On the components of the dielectric constants of ionic liquids: Ionic polarization? *Physical Chemistry Chemical Physics* 11: 2452–2458. http://dx.doi.org/10.1039/B815835E.

Jaafar MZ, Vinogradov J, and Jackson MD (2009) Measurement of streaming potential coupling coefficient in sandstones saturated with high salinity NaCl brine. *Geophysical Research Letters* 36(21).

Jackson MD (2010) Multiphase electrokinetic coupling: Insights into the impact of fluid and charge distribution at the pore scale from a bundle of capillary tubes model. *Journal of Geophysical Research* 115: B07206. http://dx.doi.org/10.1029/2009JB007092.

Jackson MD, Gulamali MY, Leinov E, Saunders JH, and Vinogradov J (2010) Real-time measurements of spontaneous potential for inflow monitoring in intelligent wells. *Proceedings – SPE Annual Technical Conference and Exhibition* 2010, pp. 4133–4150.

Jackson PD, Northmore KJ, Meldrum PI, et al. (2002) Non-invasive moisture monitoring within an earth embankment – A precursor to failure. *NDT and E International* 35(2): 107–115.

Jha SN, Narsaiah K, Basediya AL, et al. (2011) Measurement techniques and application of electrical properties for nondestructive quality evaluation of foods—A review. *Journal of Food Science and Technology* 48(4): 387–411.

Jindal M, Kumar V, Rana V, and Tiwary AK (2013) Physico-chemical, mechanical and electrical performance of bael fruit gum-chitosan IPN films. *Food Hydrocolloids* 30(1): 192–199.

Johnson DL, Koplik J, and Schwartz LM (1986) New pore-size parameter characterizing transport in porous media. *Physical Review Letters* 57(20): 2564–2567. http://dx.doi.org/10.1103/PhysRevLett.57.2564.

Jones SB and Or D (2002) Surface area, geometrical and configurational effects on permittivity of porous media. *Journal of Non-Crystalline Solids* 305(1–3): 247–254. http://dx.doi.org/10.1016/S0022-3093(02)01098-0.

Jouniaux L and Bordes C (2012) Frequency-dependent streaming potentials: A review. *International Journal of Geophysics* 11. http://dx.doi.org/10.1155/2012/648781, Article ID 648781.

Jouniaux L and Ishido T (2012) Electrokinetics in earth sciences: A tutorial. *International Journal of Geophysics* http://dx.doi.org/10.1155/2012/286107. 16 pp., Article ID: 286107.

Juhasz I (1981) Normalised Qv-the key to shaly sand evaluation using the Waxman-Smits equation in the absence of core data. In: *Society of Professional Well Log Analysts 22nd Annual Logging Symposium Transactions, Paper Z*, 36 pp.

Kanoun O (2012) *Lecture Notes on Impedance Spectroscopy: Measurement, Modeling and Application*. CRC Press, 112 pp.

Keller GV (1989) Electrical properties. In: Carmichael RS (ed.) *Practical Handbook of Physical Properties of Rocks*. Boca Raton: CRC Press, ISBN: 9780849337031.

Keller GV and Frischknecht FC (1966) *Electrical Methods in Geophysical Prospecting*. Pergamon Press, 517 pp.

Kemna A, Binley A, Cassiani G, et al. (2012) An overview of the spectral induced polarization method for near-surface applications. *Near Surface Geophysics* 10(6): 453–468. http://dx.doi.org/10.3997/1873-0604.2012027.

Key K, Constable S, LIU L, and Pommier A (2013) Electrical image of passive mantle upwelling beneath the northern East Pacific Rise. *Nature* 495(7442): 499–502.

Kirkpatrick S (1973) Hopping conduction: Experiment versus theory. In: Stuke J and Brenig W (eds.) *5th Annual Conference on Amorphous and Liquid Semiconductors: Proceedings*, pp. 103–107. Philadelphia, PA: Taylor and Francis.

Knight R and Abad A (1995) Rock/water interaction in dielectric properties: Experiments with hydrophobic sandstones. *Geophysics* 60(2): 431–436.

Knight R and Endres A (1990) A new concept in modeling the dielectric response of sandstones: Defining a wetted rock and bulk water system. *Geophysics* 55(5): 586–594.

Knight R and Nur A (1984) Effect of level of water saturation on the dielectric constant of sandstones. In: *Transactions of the SPWLA Annual Logging Symposium (Society of Professional Well Log Analysts)*.

Knight RJ and Nur A (1987a) Geometrical effects in the dielectric response of partially saturated sandstones. *The Log Analyst* 28(6): 513–519.

Knight RJ and Nur A (1987b) Dielectric constant of sandstones, 60 kHz to 4 MHz. *Geophysics* 52(5): 644–654.

Knight RJ, Nur A, and Raistrick ID (1985) Modelling the electrical response of sandstones with an equivalent circuit. In: *Transactions of the SPWLA Annual Logging Symposium (Society of Professional Well Log Analysts)*.

Korvin G (1982) Axiomatic characterization of the general mixture rule. *Geoexploration* 19(4): 267–276. http://dx.doi.org/10.1016/0016-7142(82)90031-X.

Kosmulski M (1996) Adsorption of cadmium on alumina and silica: Analysis of the values of stability constants of surface complexes calculated for different parameters of triple layer model. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 117: 201–214. http://dx.doi.org/10.1016/0927-7757(96)03706-5.

Kruschwitz S (2007) *Assessment of the complex resistivity behavior of salt affected building materials*. Diss. Fakultät VI, Planen Bauen Umwelt. Technischen Univ. Berlin, Berlin.

Kruschwitz S, Binley A, Lesmes D, and Elshenawy A (2010) Textural controls on low frequency electrical spectra of porous media. *Geophysics* 75(4): WA113–WA123. http://dx.doi.org/10.1190/1.3479835.

Lasia A (1999) In: Conway BE, Bockris J, and White RE (eds.) *Electrochemical Impedance Spectroscopy and Its Applications, Modern Aspects of Electrochemistry*, vol. 32, pp. 143–248. New York: Kluwer Academic/Plenum Publishers.

Le Méhauté A and Crepy G (1983) Introduction to transfer and transport in fractal media. *Solid State Ionics* 9&10: 17–30.

Leinov E, Vinogradov J, and Jackson MD (2010) Salinity dependence of the thermoelectric coupling coefficient in brine-saturated sandstones. *Geophysical Research Letters* 37(23).

Leroy P and Revil A (2004) A triple-layer model of the surface electrochemical properties of clay minerals. *Journal of Colloid and Interface Science* 270: 371–380. http://dx.doi.org/10.1016/j.jcis.2003.08.007.

Leroy P and Revil A (2009) A mechanistic model for the spectral induced polarization of clay materials. *Journal of Geophysical Research* 114: B10202. http://dx.doi.org/10.1029/2008JB006114.

Leroy P, Revil A, Kemna A, Cozensa P, and Ghorbani A (2008) Complex conductivity of water-saturated packs of glass beads. *Journal of Colloid and Interface Science* 321: 103–117. http://dx.doi.org/10.1016/j.jcis.2007.12.031.

Lesparre N, Gibert D, Nicollin F, Nussbaum C, and Adler A (2013) Monitoring the excavation damaged zone by three-dimensional reconstruction of electrical resistivity. *Geophysical Journal International* 195(2): 972–984.

Li D, Jiang J, Wu J, Zhang D, and Zhang X (2005) Experimental research and statistical analysis on the dielectric properties of lunar soil simulators. *Chinese Science Bulletin* 50(10): 1034–1043.

Li SX, Pengra DB, and Wong PZ (1995) Onsager's reciprocal relation and the hydraulic permeability of porous media. *Physical Review E* 51: 5748–5751. http://dx.doi.org/10.1103/PhysRevE.51.5748.

Li C, Tercier P, and Knight R (2001) Effect of sorbed oil on the dielectric properties of sand and clay. *Water Resources Research* 37(6): 1783–1793. http://dx.doi.org/10.1029/2001WR900006.

Lichtenecker K and Rother K (1931) 'Die Herleitung des logarithmischen Mischungs-gesetzes aus allegemeinen Prinzipien der stationaren Stromung. *Physikalishce Zeitschrift* 32: 255–260.

Lide DR (2012) *Handbook of Chemistry and Physics*, 93rd edn. Taylor Francis, ISBN: 1439880492.

Liu LP (2010) Hashin-Shtrikman bounds and their attainability for multi-phase composites. *Proceedings of the Royal Society A* 466: 3693–3713.

Lockner DA and Byerlee JD (1985) Complex resistivity measurements of confined rocks. *Journal of Geophysical Research* 90: 78377847.

Lorne B, Perrier F, and Avouac J-P (1999a) Streaming potential measurements: 1. Properties of the electrical double layer from crushed rock samples. *Journal of Geophysical Research, B: Solid Earth* 104(B8): 17857–17877.

Lorne B, Perrier F, and Avouac J-P (1999b) Streaming potential measurements 2. Relationship between electrical and hydraulic flow patterns from rock samples during deformation. *Journal of Geophysical Research, B: Solid Earth* 104(8): 17879–17896.

Luo M, Wood JR, and Cathles LM (1994) Prediction of thermal conductivity in reservoir rocks using fabric theory. *Journal of Applied Geophysics* 32(4): 321–334. http://dx.doi.org/10.1016/0926-9851_94_90031-0.

Macdonald JR and Kenan WR (1987) *Impedance Spectroscopy: Emphasizing Solid Materials and Systems*. John Wiley & Sons, 368 pp.

Maineult A, Bernabé Y, and Ackerer P (2004) Electrical response of flow, diffusion, and advection in a laboratory sand box. *Vadose Zone Journal* 3(4): 1180–1192.

Maineult A, Bernabé Y, and Ackerer P (2005) Detection of advected concentration and pH fronts from self-potential measurements. *Journal of Geophysical Research, B: Solid Earth* 110(11): 1–14.

Malmberg C and Maryott A (1956) Dielectric constant of water from 0 °C to 100 °C. *Journal of Research of the National Bureau of Standards* 56(1): 1–8, Paper 2641.

Marshall DJ and Madden TK (1959) Induced polarization, a study of its causes. *Geophysics* 24: 790–816.

Maurício AM, Pacheco AMG, Brito PSD, Castro B, Figueiredo C, and Aires-Barros L (2005) An ionic conductivity-based methodology for monitoring salt system in monument stones. *Journal of Cultural Heritage* 6(4): 287–293.

Maxwell JC (1873) *A Treatise on Electricity and Magnetism*. vol. 1. Oxford, UK: Clarendon Press.

Mendelson KS and Cohen MH (1982) The effect of grain anisotropy on the electrical properties of sedimentary rocks. *Geophysics* 47(2): 257–263.

Mewafy FM, Werkema DD, Atekwana EA, et al. (2013) Evidence that bio-metallic mineral precipitation enhances the complex conductivity response at a hydrocarbon contaminated site. *Journal of Applied Geophysics* 98: 113–123. http://dx.doi.org/10.1016/j.jappgeo.2013.08.011.

Mizutani H, Ishido T, Yokokura T, and Ohnishi S (1976) Electrokinetic phenomena associated with earthquakes. *Geophysical Research Letters* 3(7): 365–368.

Montaron B (2009) Connectivity theory – A new approach to modelling non-Archie rocks. *Petrophysics* 50: 102–115.

Montaron B and Han M (2009) A connectivity model for the electrical conductivity of sandstone rocks. In: *SPWLA 50th Annual Logging Symposium, June 21–24*, pp. 1–11.

Morgan FD, Williams ER, and Madden TR (1989) Streaming potential properties of westerly granite with applications. *Journal of Geophysical Research* 94: 12449–12461. http://dx.doi.org/10.1029/JB094iB09p12449.

Nugent WH, Coates GR, and Peebler RP (1978) New approach to carbonate analysis. In: *O-SPWLA Conference Paper*.

Nurmi RD and Frisinger MR (1983) *Synergy of core petrophysical measurements, log data, and rock examination in carbonate reservoir*. Society of Petroleum Engineers of AIME, (Paper) SPE 1983.

Oda M, Katsube T, and Takemura T (2002) Microcrack evolution and brittle failure of Inada granite in triaxial compression tests at 140 MPa. *Journal of Geophysical Research* 107(B10): 2233. http://dx.doi.org/10.1029/2001JB000272.

Olhoeft GR (1981) Electrical properties of rocks. In: Touloukian YS, Judd WR, and Roy RF (eds.) *Physical Properties of Rocks and Minerals*. New York: McGraw Hill.

Oliver DP, Bramley RGV, Riches D, Porter I, and Edwards J (2013) Review: Soil physical and chemical properties as indicators of soil quality in Australian viticulture. *Australian Journal of Grape and Wine Research* 19(2): 129–139.

Orazem ME and Tribollet B (2011) *Electrochemical Impedance Spectroscopy*. John Wiley & Sons, 560 pp.

Overbeek JTG (1952) Electrochemistry of the double layer. In: Kruyt HR (ed.) *Colloid Science 1, Irreversible Systems*, pp. 115–193. Elsevier.

Packard RG (1953) Streaming potentials across glass capillaries for sinusoidal pressure. *Journal of Chemical Physics* 21: 303–307.

Palacky GV (1987) Resistivity characteristics of geologic targets. *Electromagnetic Methods in Applied Geophysics, 1. Theory*, vol. 1351.

Parkhomenko EJ (1965) *Elektriceskie Svoitsva Gornich Porod*. Nauka, Moskva: Izdat.

Parks G (1965) The isoelectric points of solid oxides, solid hydroxides, and aqueous hydroxocomplex. *Chemical Reviews* 65: 177–198.

Partzsch GM, Schilling FR, and Arndt J (2000) The influence of partial melting on the electrical behavior of crustal rocks: Laboratory examinations, model calculations and geological interpretations. *Tectonophysics* 317(3–4): 189–203. http://dx.doi.org/10.1016/S0040-1951(99)00320-0.

Patchett JG (1975) An investigation of shale conductivity. In: *Paper Presented at SPWLA 16th Annual Logging Symposium, Paper V, SPWLA, New Orleans, LA, 4–7 June*, 40 pp.

Pearce JA and Zuluaga A (2004) Electrical properties of soils at ISM radio frequencies, from 1 to 40 MHz. *Journal of Microwave Power and Electromagnetic Energy* 39(3–4): 179–190.

Pengra DB, Li SX, and Wong PZ (1999) Determination of rock properties by low-frequency AC electrokinetics. *Journal of Geophysical Research* 104: 29485–29508. http://dx.doi.org/10.1029/1999JB900277.

Peplinski NR, Ulaby FT, and Dobson MC (1995) Dielectric properties of soils in the 0.3–1.3-GHz range. *IEEE Transactions on Geoscience and Remote Sensing* 33(3).

Personna YR, Ntarlagiannis D, Slater L, Yee N, O'Brien M, and Hubbard S (2008) Spectral induced polarization and electrodic potential monitoring of microbially mediated iron sulfide transformations. *Journal of Geophysical Research: Biogeosciences* 113(2): G02020. http://dx.doi.org/10.1029/2007JG000614.

Personna YR, Slater L, Ntarlagiannis D, Werkema D, and Szabo Z (2013a) Complex resistivity signatures of ethanol in sand-clay mixtures. *Journal of Contaminant Hydrology* 149: 76–87. http://dx.doi.org/10.1016/j.jconhyd.2013.07.005.

Personna YR, Slater L, Ntarlagiannis D, Werkema D, and Szabo Z (2013b) Electrical signatures of ethanol–liquid mixtures: Implications for monitoring biofuels migration in the subsurface. *Journal of Contaminant Hydrology* 144: 99–107. http://dx.doi.org/10.1016/j.jconhyd.2012.10.011.

Phillips SL, Ozbek H, and Otto RJ (1978) Basic energy properties of electrolytic solutions database. In: *Sixth International CODATA Conference Santa Flavia (Palermo), Sicily, Italy, May 22–25*. http://www.osti.gov/bridge/purl.cover.jsp;jsessionid=3954E775156A8BC0FA35DB5CE5B402D4?purl=/6269880-iPJPhB/ accessed 10 June 2010.

Pierson HO (2004) *Handbook of Carbon, Graphite, Diamond and Fullerenes – Properties, Processing and Applications*. William Andrew Publishing/Noyes.

Placencia-Gómez E, Slater L, Ntarlagiannis D, and Binley A (2013) Laboratory SIP signatures associated with oxidation of disseminated metal sulfides. *Journal of Contaminant Hydrology* 148: 25–38. http://dx.doi.org/10.1016/j.jconhyd.2013.02.007.

Poirier J-P (2000) *Introduction to the Physics of the Earth's Interior*, 2nd edn Cambridge University Press, ISBN: 9780521663922, 328 pp.

Poupon A and Leveaux J (1971) Evaluation of water saturation in shaly formations. In: *Society of Professional Well Log Analysts 12th Annual Logging Symposium Transactions, Paper O*, 2 pp.

Poupon A, Loy ME, and Tixier MP (1954) A contribution to electric log interpretation in shaly sands. *Transactions of the American Institute of Mechanical Engineers* 201: 138–145.

Pride S (1994) Governing equations for the coupled electromagnetics and acoustics of porous media. *Physical Review B* 50: 15678–15696. http://dx.doi.org/10.1103/PhysRevB.50.15678.

Pride SR and Morgan FD (1991) Electro-kinetic dissipation induced by seismic waves. *Geophysics* 56: 914–925. http://dx.doi.org/10.1190/1.1443125.

Quincke GH (1859) Ueber eine neue Art elektrischer Ströme. *Annalen de Physik und Chemie* 183(5): 1–47. http://dx.doi.org/10.1002/andp.18591830502.

Rasmus JC (1983) Variable cementation exponent, m, for fractured carbonates. *The Log Analyst* 24(6): 13–23.

Reppert PM (2000) Electrokinetics in the Earth. PhD Thesis, Massachusetts Institute of Technology.

Reppert PM and Morgan FD (2001) Streaming potential collection and data processing techniques. *Journal of Colloid and Interface Science* 233(2): 348–355.

Reppert PM and Morgan FD (2003a) Temperature-dependent streaming potentials: 1. Theory. *Journal of Geophysical Research, B: Solid Earth* 108(11), ECV3-1–ECV3-12.

Reppert PM and Morgan FD (2003b) Temperature-dependent streaming potentials: 2. Laboratory. *Journal of Geophysical Research, B: Solid Earth* 108(11), ECV4-1–ECV4-13.

Reppert PM, Morgan FD, Lesmes DP, and Jouniaux L (2001) Frequency-dependent streaming potentials. *Journal of Colloid and Interface Science* 234: 194–203.

Reuβ FF (1809) Sur un nouvel effet de l'électricité galvanique. *Mémoires de la Société Impériale des Naturalistes de Moskou* 2: 327–337.

Revil A and Cerepi A (2004) Streaming potentials in two-phase flow conditions. *Geophysical Research Letters* 31: L11605. http://dx.doi.org/10.1029/2004GL020140.

Revil A and Florsch N (2010) Determination of permeability from spectral induced polarization in granular media. *Geophysical Journal International* 181(3): 1480–1498.

Revil A and Glover PWJ (1997) Theory of ionic-surface electrical conduction in porous media. *Physical Review B – Condensed Matter and Materials Physics* 55(3): 1757–1773.

Revil A and Glover PWJ (1998) Nature of surface electrical conductivity in natural sands, sandstones, and clays. *Geophysical Research Letters* 25(5): 691–694.

Revil A and Leroy P (2001) Hydroelectric coupling in a clayey material. *Geophysical Research Letters* 28(8): 1643–1646. http://dx.doi.org/10.1029/2000GL012268.

Revil A and Leroy P (2004) Constitutive equations for ionic transport in porous shales. *Journal of Geophysical Research* 109: B03208. http://dx.doi.org/10.1029/2003JB002755.

Revil A and Skold M (2011) Salinity dependence of spectral induced polarization in sands and sandstones. *Geophysical Journal International* 187(2): 813–824.

Revil A, Cathles LM, Losh S, and Nunn JA (1998) Electrical conductivity in shaly sands with geophysical applications. *Journal of Geophysical Research, B: Solid Earth* 103(10): 23925–23936.

Revil A, Pezard PA, and Glover PWJ (1999a) Streaming potential in porous media 1. Theory of the zeta potential. *Journal of Geophysical Research, B: Solid Earth* 104(B9): 20021–20031.

Revil A, Schwaeger H, Cathles LM III, and Manhardt PD (1999b) Streaming potential in porous media. 2. Theory and application to geothermal systems. *Journal of Geophysical Research, B: Solid Earth* 104(B9): 20033–20048.

Revil A, Hermitte D, Spangenberg E, and Cochemé JJ (2002) Electrical properties of zeolitized volcaniclastic materials. *Journal of Geophysical Research, B: Solid Earth* 107(8), ECV3-1–ECV3-17.

Revil A, Linde N, Cerepi A, Jougnot D, Matthäi S, and Finsterle S (2007) Electrokinetic coupling in unsaturated porous media. *Journal of Colloid and Interface Science* 313(1): 315–327.

Revil A, Johnson TC, and Finizola A (2010) Three-dimensional resistivity tomography of Vulcan's forge, Vulcano Island, southern Italy. *Geophysical Research Letters* 37(15), L15308.

Reynolds JA and Hough JM (1957) Formulae for dielectric constant of mixtures. *Proceedings of the Physical Society* 70: 769–775.

Robinson DA, Jones SB, Wraith JM, Ord D, and Friedman SP (2003) A review of advances in dielectric and electrical conductivity measurement in soils using time domain reflectometry. *Vadose Zone Journal* 2(4): 444–475. http://dx.doi.org/10.2136/vzj2003.4440.

Rojo A, Hansen HK, and Monárdez O (2014) Electrokinetic remediation of mine tailings by applying a pulsed variable electric field. *Minerals Engineering* 55: 52–56.

Ross DS and Ketterings Q (2011) *Recommended methods for determining soil cation exchange capacity. Recommended Soil Testing Procedures for the Northeastern United States*. pp. 75–85, Northeastern Regional Publication no. 493.

Ruffet C, Gueguen Y, and Darot M (1991a) Complex conductivity measurements and fractal nature of porosity. *Geophysics* 56(6): 758–768. http://dx.doi.org/10.1190/1.1443093.

Ruffet C, Gueguen Y, and Darot M (1991b) Rock conductivity and fractal nature of porosity. *Terra Nova* 3(3): 265–275. http://dx.doi.org/10.1111/j.1365-3121.1991.tb00144.x.

Sass O and Viles HA (2010) Wetting and drying of masonry walls: 2D-resistivity monitoring of driving rain experiments on historic stonework in Oxford, UK. *Journal of Applied Geophysics* 70(1): 72–83.

Saunders JH, Jackson MD, Gulamali MY, Vinogradov J, and Pain CC (2012) Streaming potentials at hydrocarbon reservoir conditions. *Geophysics* 77(1): E77–E90.

Saunders JH, Jackson MD, and Pain CC (2008) Fluid flow monitoring in oilfields using downhole measurements of electrokinetic potential. *Geophysics* 73(5): E165–E180. http://dx.doi.org/10.1190/1.2959139.

Scales PJ, Grieser F, and Healy TW (1990) Electrokinetics of the muscovite mica-aqueous interface. *Langmuir* 6: 582–589. http://dx.doi.org/10.1021/la00093a012.

Schilling F, Partzsch GM, Brasse H, and Schwartz G (1997) Partial melting below the magmatic arc in the central Andes deduced from geoelectric field experiments and laboratory data. *Physics of the Earth and Planetary Interiors* 103: 17–31.

Schlumberger (1972) *Log Interpretation, Volume 1 – Principles*. New York: Schlumberger Ltd., 113 pp.

Schlumberger (2013) Log Interpretation Charts, Gen-4 (Gen-8 in previous annually issued charts), Schlumberger.

Schmutz M, Revil A, Vaudelet P, Batzle M, Viñao PF, and Werkema DD (2010) Influence of oil saturation upon spectral induced polarization of oil-bearing sands. *Geophysical Journal International* 183(1): 211–224.

Schön JH (2004) Physical Properties of Rocks: Fundamentals and Principles of Petrophysics. In: Helbig K and Treitel S (eds.) *vol. 18* Amsterdam, The Netherlands: Elsevier, ISBN: 008044346X.

Schwartz LM, Sen PN, and Johnson DL (1989) Influence of rough surfaces on electrolytic conduction in porous media. *Physical Review B* 40(4): 2450–2458. http://dx.doi.org/10.1103/PhysRevB.40.2450.

Schwarz G (1962) A theory of the low-frequency dielectric dispersion of colloidal particles in electrolyte solution. *Journal of Physical Chemistry* 66: 2636–2642.

Scott JBT and Barker RD (2003) Determining pore-throat size in Permo-Triassic sandstones from low-frequency electrical spectroscopy. *Geophysical Research Letters* 30(9): 1450. http://dx.doi.org/10.1029/2003GL016951.

Scott JBT and Barker RD (2005) Characterization of sandstone by electrical spectroscopy for stratigraphical and hydrogeological investigations. *Quarterly Journal of Engineering Geology & Hydrogeology* 38: 143–154.

Seager CH and Pike GE (1974) Percolation and conductivity: A computer study II. *Physical Review B* 10: 1435–1446.

Sears AR and Groves JN (1978) The use of oscillating laminar flow streaming potential measurements to determine the zeta potential of a capillary. *Journal of Colloid and Interface Science* 65: 479–482.

Selway K (2014) On the Causes of Electrical Conductivity Anomalies in Tectonically Stable Lithosphere. *Surveys in Geophysics* 35(1): 219–257.

Sen P and Goode P (1992a) Influence of temperature on electrical conductivity on shaly sands. *Geophysics* 57: 89–96.

Sen P and Goode P (1992b) Errata, to: Influence of temperature on electrical conductivity of shaly sands. *Geophysics* 57: 1658.

Sen C, Scala PN, and Cohen MH (1981) A self-similar model for sedimentary rocks with application to the dielectric constant of fused glass beads. *Geophysics* 46(5): 781–795.

Shah C and Yortsos Y (1996) The permeability of strongly disordered systems. *Physics of Fluids* 8: 280–282.

Sheffer MR, Reppert PM, and Howie JA (2007) A laboratory apparatus for streaming potential and resistivity measurements on soil samples. *Review of Scientific Instruments* 78(9): 094502. http://dx.doi.org/10.1063/1.2782710.

Simandoux P (1963) Dielectric measurements on porous media application to the measurement of water saturations: Study of the behavior of argillaceous formations (Mesures dielectriques en milieu poreux, application a la Mewe des saturations en eau, etude du comportement des massifs argileux). *Revue de l'Institut Francais du Petrole* 18: 193–215, Supplementary Issue.

Skold M, Revil A, and Vaudelet P (2011) The pH dependence of spectral induced polarization of silica sands: Experiment and modelling. *Geophysical Research Letters* 38(12), L12304.

Slater L (2007) Near surface electrical characterization of hydraulic conductivity: From petrophysical properties to aquifer geometries—A review. *Surveys in Geophysics* 28: 169–197. http://dx.doi.org/10.1007/s10712-007-9022-y.

Slater L and Glaser D (2003) Controls on induced polarization in sandy unconsolidated sediments and application to aquifer characterization. *Geophysics* 68(5): 1547–1558.

Slater L and Lesmes DP (2002) Electrical-hydraulic relationships observed for unconsolidated sediments. *Water Resources Research* 38(10): 1213.

Slater L, Ntarlagiannis D, and Wishart D (2006) On the relationship between induced polarization and surface area in metal-sand and clay-sand mixtures. *Geophysics* 71(2): A1–A5.

Smith WDM and Rouleau RJ (1977) A modified shaly sand Sw equation for the Western Canada Cretaceous. In: *Canadian Well Logging Society Sixth Formation Evaluation Conference, Paper V*, 4 pp.

Smoluchowski M (1903) Contribution à la théorie l´endosmose électrique et de quelques phénomènes corrélatifs. *Krak. Anz.* 1903: 182–199.

Somerton WH (1992) *Thermal Properties and Temperature-Related Behavior of Rock/Fluid Systems*. Amsterdam, The Netherlands: Elsevier, ISBN: 0444890017.

Speight JG (2011) *An Introduction to Petroleum Technology, Economics, and Politics*. John Wiley & Sons978-1-118-19254-2, 334 pp.

Sposito G (1989) *The Chemistry of Soils*. Oxford: Oxford University.

Sprunt ES, Mercer TB, and Djabbarah NF (1994) Streaming potential from multiphase flow. *Geophysics* 59: 707–711. http://dx.doi.org/10.1190/1.1443628.

Susanto A, Koleva DA, Copuroglu O, Van Beek K, and Van Breugel K (2013) Mechanical, electrical and microstructural properties of cement-based materials in conditions of stray current flow. *Journal of Advanced Concrete Technology* 11(3): 119–134.

Sweeney SA and Jennings HY Jr. (1960) Effect of wettability on the electrical resistivity of carbonate rock from a petroleum reservoir. *Journal of Physical Chemistry* 64(5): 551–553. http://dx.doi.org/10.1021/j100834a009.

Tardif E, Glover PWJ, and Ruel J (2011) Frequency-dependent streaming potential of Ottawa sand. *Journal of Geophysical Research B* 116: B04206. http://dx.doi.org/10.1029/2010JB008053.

Thompson A and Gist G (1993) Geophysical applications of electrokinetic conversion. *The Leading Edge* 12: 1169–1173.

Thomson WT and Dahleh MD (1998) *Theory of Vibration with Applications*, 5th edn. Englewood Cliffs, NJ: Prentice-Hall, 524 pp.

Tiab D and Donaldson EC (2011) *Petrophysics, Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties*, 3rd edn. Gulf Professional Publishing, ISBN: 9780123838483.

Titov K, Komarov V, Tarasov V, and Levitski A (2002) Theoretical and experimental study of time domain-induced polarization in water-saturated sands. *Journal of Applied Geophysics* 50: 417–433.

Titov K, Tarasov A, Yuri I, Seleznev N, and Boyd A (2010) Relationships between induced polarization relaxation time and hydraulic properties of sandstone. *Geophysical Journal International* 180: 1095–1106.

Ulrich C and Slater L (2004) Induced polarization measurements on unsaturated, unconsolidated sands. *Geophysics* 69(3): 762–771.

Vanhala H (1997) Mapping oil-contaminated sand and till with the spectral induced polarization (SIP) method. *Geophysical Prospecting* 45(2): 303–326.

Vaudelet P, Revil A, Schmutz M, Franceschi M, and Bégassat P (2011a) Changes in induced polarization associated with the sorption of sodium, lead, and zinc on silica sands. *Journal of Colloid and Interface Science* 360(2): 739–752.

Vaudelet P, Revil A, Schmutz M, Franceschi M, and Bégassat P (2011b) Induced polarization signatures of cations exhibiting differential sorption behaviors in saturated sands. *Water Resources Research* 47(2), W02526.

Vinogradov J and Jackson MD (2011) Multiphase streaming potential in sandstones saturated with gas/brine and oil/brine during drainage and imbibition. *Geophysical Research Letters* 38: L01301. http://dx.doi.org/10.1029/2010GL045726.

Vinogradov J, Jaafar MZ, and Jackson MD (2010) Measurement of streaming potential coupling coefficient in sandstones saturated with natural and artificial brines at high salinity. *Journal of Geophysical Research, B: Solid Earth* 115(12).

Volkmann J and Klitzsch N (2010) Frequency-dependent electric properties of microscale rock models for frequencies from one millihertz to ten kilohertz. *Vadose Zone Journal* 9: 858–870. http://dx.doi.org/10.2136/vzj2009.0162.

Waff HS (1974) Theoretical consideration of electrical conductivity in a partially molten mantle and implications for geothermometry. *Journal of Geophysical Research* 79: 4003–4010.

Wagner KW (1914) Erklärung der dielektrischen Nachwirkunsvorgänge auf Grund Maxwellscher Vorstellungen. *Archiv fur Elektrotechnik* 2(9): 371–387.

Walker E and Glover PWJ (2010) Characteristic pore size, permeability and the electro-kinetic coupling coefficient transition frequency in porous media. *Geophysics* 75(6): E235–E246.

Walker E, Glover PWJ, and Ruel J (2014) A transient method for measuring the DC streaming potential coefficient of porous and fractured rocks. *Journal of Geophysical Research - Solid Earth* 119. http://dx.doi.org/10.1002/2013JB010579.

Wall S (2010) The history of electrokinetic phenomena. *Current Opinion in Colloid and Interface Science* 15: 119–124. http://dx.doi.org/10.1016/j.cocis.2009.12.005.

Waxman MH and Smits LJM (1968) Electrical conductivities in oil-bearing shaly sands. *SPE Journal* 8(2): 107–122. http://dx.doi.org/10.2118/1863-A.

Weiss J, Snyder K, Bullard J, and Bentz D (2013) Using a saturation function to interpret the electrical properties of partially saturated concrete. *Journal of Materials in Civil Engineering* 25(8): 1097–1106.

Weller A, Breede K, Slater L, and Nordsiek S (2011) Effect of changing water salinity on complex conductivity spectra of sandstones. *Geophysics* 76(5): F315–F327. http://dx.doi.org/10.1190/geo2011-0072.1.

Winsauer WO, Shearin HM, Masson PH, and Williams M (1952) Resistivity of brine-saturated sands in relation to pore geometry. *AAPG Bulletin* 36: 253–277.

Wittmann K, Feuerbacher B, Ulamec S, et al. (1999) Rosetta lander in situ characterization of a comet nucleus. *Acta Astronautica* 45(4–9): 389–395.

Wong PZ (1987) Fractal surface in porous media. In: Banavar JR, Koplik J, and Winkler KW (eds.) *Physics and Chemistry of Porous Media,* vol. 2, *Am. Inst. Phys. Conference Proceedings*, *vol. 154*, pp. 304–318.

Worthington PF (1985) The evolution of shaly-sand concepts in reservoir evaluation. *The Log Analyst* XXVI(1): 23–40.

Wu L, Forsling W, and Schlindler PW (1991) Surface complexation of calcium minerals in aqueous solution. *Journal of Colloid and Interface Science* 147: 178–185. http://dx.doi.org/10.1016/0021-9797(91)90145-X.

Wyllie MRJ and Gregory AR (1953) Formation factors of unconsolidated porous media: Influence of particle shape and effect of cementation. *Transactions of the AIME* 198: 103–110.

Xu Y and Schoonen MAA (2000) The absolute energy positions of conduction and valence bands of selected semiconducting minerals. *American Mineralogist* 85(3–4): 543–556.

Xu Y, Shankland TJ, and Duba AG (2000) Pressure effect on electrical conductivity of mantle olivine. *Physics of the Earth and Planetary Interiors* 118: 149.

Yang X (2011) Origin of high electrical conductivity in the lower continental crust: A review. *Surveys in Geophysics* 32(6): 875–903.

Yoshino T and Noritake F (2011) Unstable graphite films on grain boundaries in crustal rocks. *Earth and Planetary Science Letters* 306(3–4): 186–192.

Yoshino T, Nishi M, Matsuzaki T, Yamazaki D, and Katsura T (2008) Electrical conductivity of majorite garnet and its implications for electrical structure in the mantle transition zone. *Physics of the Earth and Planetary Interiors* 170(3–4): 193–200.

Zakri T, Laurent J-P, and Vauclin M (1998) Theoretical evidence for 'Lichtenecker's mixture formulae' based on the effective medium theory. *Journal of Physics D: Applied Physics* 31: 1589–1594.

Zhang J and Catchmark JM (2011) A catalytically powered electro-kinetic lens, Toward channel-less microfluidics. *Microfluidics and Nanofluidics* 10(5): 1147–1151.

Zhang C, Ntarlagiannis D, Slater L, and Doherty R (2010) Monitoring microbial sulfate reduction in porous media using multipurpose electrodes. *Journal of Geophysical Research, Biogeosciences* 15(G3). http://dx.doi.org/10.1029/2009JG001157.

Zhang C, Slater L, and Prodan C (2013) Complex dielectric properties of sulfate-reducing bacteria suspensions. *Geomicrobiology Journal* 30(6): 490–496. http://dx.doi.org/10.1080/01490451.2012.719997.

Zhang C, Slater L, Redden G, Fujita Y, Johnson T, and Fox D (2012) Spectral induced polarization signatures of hydroxide adsorption and mineral precipitation in porous media. *Environmental Science & Technology* 46: 4357–4364. http://dx.doi.org/10.1021/es204404e.

Zhao JH, Rivera E, Mufti A, Stephenson D, and Thomson DJ (2012) Evaluation of dielectric based and other methods for moisture content measurement in building stones. *Journal of Civil Structural Health Monitoring* 2(3–4): 137–148.

This page intentionally left blank

11.05 Geophysical Properties of the Near-Surface Earth: Magnetic Properties

BM Moskowitz and M Jackson, University of Minnesota, Minneapolis, MN, USA
V Chandler, Minnesota Geological Survey, St. Paul, MN, USA

© 2015 Elsevier B.V. All rights reserved.

11.05.1 Introduction 141
11.05.2 Basic Concepts 141
11.05.2.1 Magnetic Quantities 141
11.05.2.1.1 Induction and magnetic field 141
11.05.2.1.2 Magnetization 142
11.05.2.1.3 Susceptibility and permeability 142
11.05.2.2 Classes of Magnetic Behavior 143
11.05.2.2.1 Diamagnetism and paramagnetism 143
11.05.2.2.2 Ferromagnetism 144
11.05.2.2.3 Antiferromagnetism 144
11.05.2.2.4 Ferrimagnetism 145
11.05.2.2.5 Canted and defect antiferromagnetism 145
11.05.2.3 Magnetic Units 145
11.05.3 Magnetic Ordering and Magnetic Minerals 146
11.05.3.1 Iron Oxides 146
11.05.3.1.1 Magnetite 146
11.05.3.1.2 Titanomagnetite 148
11.05.3.1.3 Maghemite and titanomaghemite 148
11.05.3.1.4 Hematite 148
11.05.3.1.5 Titanohematite 149
11.05.3.2 Iron Sulfides and Other Magnetic Minerals 150
11.05.4 Induced and Remanent Magnetization 151
11.05.4.1 Magnetic Hysteresis 151
11.05.4.2 Magnetic Domains 152
11.05.4.3 Magnetism in Mineral Grains 153
11.05.4.3.1 Anisotropy energy 153
11.05.4.3.2 Magnetostatic energy 154
11.05.4.3.3 Single-domain, multidomain, and pseudo-single-domain particles 154
11.05.4.3.4 Critical sizes 155
11.05.4.3.5 Superparamagnetism 156
11.05.4.4 Thermal Relaxation and Néel Theory 156
11.05.4.4.1 DC magnetization 157
11.05.4.4.2 AC susceptibility 157
11.05.4.5 Remanence Acquisition Processes 157
11.05.4.5.1 Natural remanent magnetization 157
11.05.4.5.2 Artificial remanent magnetization 160
11.05.4.6 Complex Susceptibility and Frequency Dependence of Susceptibility 161
11.05.5 Magnetic Petrology: Induced and Remanent Magnetization of Earth Materials 164
11.05.5.1 Igneous Rocks 164
11.05.5.2 Sedimentary Rocks 167
11.05.5.3 Metamorphic Rocks 167
11.05.5.4 Unconsolidated Surficial Materials 168
11.05.5.4.1 Magnetic properties of glacial deposits 168
11.05.5.4.2 Magnetic properties of soils and contaminated soils 168
11.05.5.5 Archaeological Materials 169
11.05.6 Summary 170
Acknowledgments 170
References 170

Glossary

Anisotropy energy Energy of a magnetic particle that varies with the direction of its magnetization.

Antiferromagnetism A type of magnetic ordering where two magnetic sublattices of equal magnitude are aligned antiparallel to each other resulting in zero net magnetic moment.

Blocking temperature A temperature less than the Curie temperature, below which the ratio of magnetic anisotropy energy to thermal fluctuation energy becomes large enough to stabilize remanence (intensity and direction) over geologic timescales.

Canted antiferromagnetism A variation of antiferromagnetism where the two magnetic sublattices deviate slightly from perfect antiparallelism resulting in small net spontaneous magnetization.

Coercivity The magnetic field required to reduce the induced magnetization of a material to zero after it has been exposed to a saturating field.

Coercivity of remanence The magnetic field required to reduce the saturation remanent magnetization of a material to zero after it has been exposed to a saturating field.

Complex susceptibility Susceptibility measured in an alternating magnetic field decomposed into a component that is in-phase and one that is out-of-phase (quadrature) with the external magnetic field.

Curie temperature The magnetic ordering temperature below which the spontaneous alignment of magnetic moments in a ferromagnetic or ferrimagnetic material occurs and spontaneous magnetization appears.

Defect antiferromagnetism A variation of antiferromagnetism where lattice defects or impurity cations cause an imbalance in the perfect antiparallelism resulting in small net spontaneous magnetization.

Demagnetizing field The magnetic field (H field) created inside a magnetized body by the presence of uncompensated magnetic poles, which acts in the opposite direction to the magnetization, which created it.

Diamagnetism A property of all matter related to the precession of electron orbits in an applied field, which produces a weak magnetic moment antiparallel to the field and a small negative susceptibility that is independent of temperature.

Domain Region of a particle or crystal in which the magnetization is uniform.

Domain wall The boundary between domains where the magnetization gradually rotates from the direction in one domain to its neighbor.

Domain wall energy Energy associated with the additional exchange and magnetocrystalline energies to accommodate the spin rotation from one domain to its neighbor.

Easy axis Certain directions in a crystal or particle corresponding to the minimum in anisotropy energy and along which the magnetization will align. Also referred to as an easy direction of magnetization.

Ferrimagnetism A type of magnetic ordering where two magnetic sublattices of unequal magnitudes are aligned antiparallel to each other resulting in a net spontaneous magnetization.

Ferromagnetism A type of magnetic ordering where all the atomic spin moments are aligned parallel to each other resulting in a large spontaneous magnetization.

Induced magnetization Magnetization produced by an applied field.

Koenigsberger ratio Dimensionless ratio of natural remanent magnetization (NRM) to the induced magnetization in the geomagnetic field.

Magnetic susceptibility Ratio of induced magnetization to applied field (M/H).

Magnetic viscosity Time-dependent magnetization.

Magnetization Magnetic moment per unit volume ($\mathrm{A\,m^{-1}}$) or per unit mass ($\mathrm{A\,m^2\,kg^{-1}}$).

Magnetocrystalline anisotropy Anisotropy associated with the coupling between the spin and orbital motions of electrons resulting in certain crystalline directions being easy (minimum energy) and hard axes (maximum energy) of magnetization.

Magnetostatic energy Energy of a magnetic particle stored in the magnetic pole distribution due to uncompensated surface and volume poles and in the interaction of the magnetization with external fields.

Multidomain A magnetic state of a particle containing more than one domain where magnetization changes take place mainly by domain wall motion or by nucleation/denucleation of domains.

Néel temperature The magnetic ordering temperature below which the spontaneous alignment of magnetic moments in an antiferromagnetic material occurs.

Paramagnetism A property of certain materials that have uncompensated electron spins (permanent atomic magnetic moments); their alignment with the ambient magnetic field is disrupted by random thermal perturbations, producing a positive susceptibility that varies inversely with absolute temperature.

Pseudo-single-domain Magnetic particles larger than the equilibrium single-domain size with nonuniform magnetization but exhibiting single-domain-like characteristics.

Relaxation time A characteristic time that separates stable (blocked magnetization) from unstable (unblocked or superparamagnetic) behavior and is a strong function of temperature, applied field, particle volume, and shape.

Remanent magnetization Magnetization remaining in zero magnetic field. The intensity and direction of remanent magnetization depend on the magnetic material and its magnetic history.

Saturation magnetization In high fields, the magnetization reaches a saturation value where individual magnetic moments are maximally aligned with the applied field.

Shape anisotropy Anisotropy related to the shape of a magnetic particle originating in the shape dependence of the demagnetizing field.

Single domain A magnetic state consisting of only one domain where magnetization changes mainly by moment rotation from one easy axis to another.

Spontaneous magnetization Magnetization produced by the spontaneous alignment of magnetic moments in an atomic lattice due to exchange interactions below the Curie or Néel temperature.

Superparamagnetism State in which magnetization changes are thermally activated in times much shorter than observation times, usually at high temperatures or in small grain volumes.

11.05.1 Introduction

Magnetism is associated with only a handful of elements (Fe, Ni, Co, Cr, Mn, and some rare earths), of which iron is by far the most common in geologic materials. Rock magnetism is thus, in a broad sense, a means of studying the geology, geochemistry, and geophysics of iron. This 'ferrocentric' view of Earth materials, processes, and history is focused, but by no means narrow. Iron is one of the two most abundant elements in our planet, comprising some 28–33% of its mass (e.g., Allègre et al., 2001). It is the dominant element in the core, where its high electrical conductivity and convective motions combine to generate our self-sustaining planetary magnetic field (e.g., Merrill and McFadden, 1995). Iron is also ubiquitous in the mantle, crust, and surficial materials, where it occurs in phases, such as silicates, oxides, sulfides, and oxyhydroxides, with an immense variety of magnetic characteristics, including ferromagnetism (FM) (*sensu lato*) with its ability to acquire and retain remanent magnetization in the weak geomagnetic field. It is the remarkable property of FM that makes possible our ability to reconstruct past records of lithospheric plate motions, geomagnetic field behavior, or paleoclimate change, as well as the magnetic exploration of the subsurface. The distribution of magnetic phases in the Earth's crust gives rise to the crustal magnetic field that can be measured on length scales from the local to planetary-wide and from a variety of ground-based, borehole, airborne, marine, and satellite magnetometers (e.g., Langel and Hinze, 1998; Nabighian et al., 2005; Purucker and Whaler, 2007).

Magnetic anomalies arise either from induced magnetization in the present-day geomagnetic field related to magnetic susceptibilities or from remanent magnetizations. For almost all geologic materials, magnetic measurements provide information on and insight into the physical, chemical, and even biological processes that have affected the iron phases within. Physically, a multitude of processes act over the history of a rock or sediment unit to control the magnetization states of the remanence-carrying particles, as well as the distributions of particle sizes, shapes, and orientations of the magnetic, paramagnetic, and nonmagnetic mineral grains. Chemically, iron can change its valence state easily, ensuring relatively high reactivity in inorganic redox reactions as well as exploitation by bacteria in metabolic electron-transfer reactions. Iron biomineralization occurs in connection with magnetoreception and magnetotaxis in birds, insects, and bacteria. Pedogenesis produces magnetic nanoparticles by an incompletely understood set of organic and inorganic crystal-growth and transformation processes. Magnetic iron oxides are also destroyed by reductive dissolution in many environments.

Anthropogenic processes (both past and present) can also disturb the near-subsurface environment in ways frequently detected by magnetic surveying. Cultural features associated with buried ferrous objects or structures at archaeological sites and modern urban and industrial sites, as well as land mines and unexploded ordnance (UXO), have been the subject of magnetic studies (e.g., Eppelbaum, 2011; Hannam and Dearing, 2008; Linford, 2005). Moreover, human activities can modify the magnetic properties of soils through chemical alteration or landscape reworking whereby archaeological sites are often more magnetic than surrounding undisturbed areas (e.g., Dalan, 2008; Linford, 2005).

There are several thorough recent reviews of rock magnetism and magnetic properties of Earth materials. Fundamental rock magnetism has been reviewed by Dunlop and Özdemir (2007), Harrison et al. (2007), and Harrison and Feinberg (2009); magnetic measurement techniques and instrumentation by Turner et al. (2007); and crustal magnetism and magnetic petrology by Purucker and Whaler (2007), Gee and Kent (2007), and McEnroe et al. (2009). In addition, the *Encyclopedia of Geomagnetism and Paleomagnetism* (Gubbins and Herrero-Bervera, 2007) provides short articles on many topics on rock magnetism, magnetic properties of rocks and minerals, and magnetic techniques. In addition, several excellent textbooks cover all aspects of paleomagnetism and rock magnetism and include Butler (1992), Dunlop and Özdemir (1997), Evans and Heller (2003), Lowrie (2007), and Tauxe (2010). The two texts by Butler (1992) and Tauxe (2010) are available free online. Magnetic properties in the context of magnetic surveying and magnetic anomalies have been reviewed by Grant (1985a,b), Reynolds et al. (1990), Hunt et al. (1995), Clark (1997, 1999), Linford (2005), Dalan (2008), and Hinze et al. (2013). Paleomagnetic and rock magnetic databases and publicly available software are available through The Magnetics Information Consortium (MagIC; http://earthref.org/MAGIC/).

Our objectives in this article are to present a concise overview of the essential physical and mineralogical origins of magnetism in near-surface earth materials, to summarize the key magnetic properties of common classes of geomaterials and their contributions to local magnetic fields, and to illustrate a few selected applications of magnetism in studying surficial processes. As with other geophysical methods, a better understanding of the magnetic properties responsible for a particular anomaly is essential for a more reliable geologic, archaeological, or environmental interpretation of the subsurface.

11.05.2 Basic Concepts

11.05.2.1 Magnetic Quantities

11.05.2.1.1 Induction and magnetic field

For both historical and practical reasons, there are two vector fields that are used to describe magnetic forces and energy

in relation to electrical-current or permanent-magnet field sources: magnetic induction B (also known as magnetic flux density and measured in units of tesla or webers m^{-2}) and magnetic field H (units of A m^{-1}). The distinction between B and H is important primarily in the presence of matter and at material discontinuities, as will be described in a subsequent section. In vacuum and (to an excellent approximation) in air, these two vectors are always parallel in direction and proportional in magnitude; the proportionality constant is μ_0, the permeability of free space ($4\pi \times 10^{-7}$ henry m^{-1}).

In the vicinity of a linear filamentary current I, the B and H fields form closed circular loops, in planes perpendicular to the current and concentric with it. In the cylindrical coordinate system (r, ϕ, z) where current flow is parallel to z, the B and H fields are oriented tangentially, in the ϕ direction, and their intensity varies inversely with radial distance r:

$$\vec{B} = \frac{\mu_0 I}{2\pi r}\hat{\phi} \text{ and } \vec{H} = \frac{I}{2\pi r}\hat{\phi} \quad [1]$$

When current flows in a closed circular loop, the resultant B and H fields have a geometry that, at distances large compared to the loop diameter, is that of a magnetic dipole. In spherical coordinates (r, θ, ϕ) where current flows around the loop in the ϕ direction, the magnetic dipole fields have only poloidal (radial and colatitudinal) components:

$$\vec{B} = \left(\frac{\mu_0 m}{4\pi r^3}\right)\left[2\cos\theta\hat{r} + \sin\theta\hat{\theta}\right] \text{ and}$$

$$\vec{H} = \left(\frac{m}{4\pi r^3}\right)\left[2\cos\theta\hat{r} + \sin\theta\hat{\theta}\right] \quad [2]$$

The dipole moment m is oriented perpendicular to the loop and has a magnitude equal to the product of loop area and current strength (units A m^2).

A current loop, with its corresponding magnetic dipole moment, is subject to forces and torques in an externally applied magnetic field B_0. The Zeeman energy is given by

$$E_Z = -\vec{m}\cdot\vec{B_0} = -\mu_0 mH\cos\alpha \quad [3]$$

If B_0 is uniform, there is an aligning torque

$$\tau = d(E_Z)/d\alpha = \mu_0 mH\sin\alpha = \vec{m}\times\vec{B_0} \quad [4]$$

but no net force. When B_0 is nonuniform, there is a translational force that acts on a dipole moment, proportional to the gradient of field strength:

$$\vec{F} = -\overrightarrow{\nabla(E_Z)} = \nabla\left(\vec{m}\cdot\vec{B_0}\right) = \left(\vec{m}\cdot\vec{\nabla}\right)\vec{B_0} \quad [5]$$

where the last equality holds when m is constant, independent of B_0.

11.05.2.1.2 Magnetization

All matter responds to externally applied magnetic fields, and the response alters the local B and H fields in different ways. Magnetization M is related to the other two vector fields by $\vec{B} = \mu_0\left(\vec{H} + \vec{M}\right)$; M is measured in the same units as H (A m^{-1}). Magnetization in the absence of external fields is a unique property of certain ('ferromagnetic' or 'magnetically ordered') materials that will be described in a later section.

M changes discontinuously at the boundary between a magnetized substance and air or vacuum (where $M = 0$) or more generally between two magnetized materials with differing properties. The divergence of M is equivalent to a 'magnetic surface charge density' or 'pole density,' $\sigma_m = \left(\hat{n}\cdot\vec{M}\right)$, which produces H fields in the same way that electrical charges produce E fields. These H fields also therefore have surface-normal components that change discontinuously at material boundaries. Within a magnetized body, the H field due to surface poles is generally antiparallel to and proportional to M; the proportionality constant N is known as the demagnetizing factor. Such 'demagnetizing fields' in detail depend on particle shape and magnetization geometry; for the simple case of a uniformly magnetized ellipsoid, $\vec{H_d} = -N\vec{M}$, and for more complicated configurations, this remains a useful approximation. These demagnetizing fields are an important source of anisotropy in particles of strongly magnetic materials.

The B field, unlike H, cannot be conceived as originating in a distribution of magnetic poles. Its surface-normal component is continuous across material boundaries ($\vec{\nabla}\cdot\vec{B} = 0$), and field lines always form closed loops, rather than beginning and ending at magnetic poles as they do (in the absence of free electrical currents) for H. When B originates in magnetic material, it is most readily visualized as coming from 'bound currents' (K_m) flowing along the surface of the magnetized body $K_m = \vec{M}\times\hat{n}$.

In the space surrounding a magnetized body, the B and H fields due to M are superposed onto the externally applied field, if any. The fields due to M are dominantly dipolar (eqn [2]) at distances that are large in comparison to the size of the body, and the dipole moment $\vec{m}$ is equal to the volume integral of $\vec{M}$. In laboratory measurements, it is common to normalize measured magnetic moments by specimen mass, which is often easier to measure accurately than specimen volume; however, for field-anomaly modeling, volumetric magnetization is the quantity of interest. In older literature, mass-specific magnetization (A m^2 kg^{-1}) was often indicated by the symbol J to distinguish it from volumetric magnetization, but in modern usage, M is most commonly used for either quantity.

11.05.2.1.3 Susceptibility and permeability

The intensity of magnetization M_i induced in a volume of material by exposure to a weak external applied field B_0 or H_0 is generally proportional to the intensity of the applied field and approximately parallel to the field. Susceptibility k or χ is a symmetrical second-rank tensor defined by $\vec{M_i} = [k]\vec{H_0}$. For isotropic materials, susceptibility is the scalar ratio M_i/H_0, and even for anisotropic materials, a scalar approximation is sufficient for many purposes. When M_i is not linearly proportional to H_0, initial susceptibility is defined as $k = \lim_{H_0\to 0} dM_i/dH_0$. In an applied AC field, induced magnetization may lag behind the field due to various physical effects, and AC susceptibility is therefore a complex quantity with real (in-phase, k') and imaginary (out-of-phase, k'') components (see **Section 11.05.4.4.2**). Some authors reserve χ for mass-specific susceptibility (m^3 kg^{-1}) and k for dimensionless volumetric susceptibility (e.g., Evans and Heller, 2003; Hunt et al., 1995), and others use the opposite convention (e.g., Dunlop and Özdemir, 1997; Tauxe, 2010). In this article,

we will use χ for mass-specific susceptibility. Even though k is dimensionless, it is common to attach SI to the value to avoid confusion.

Susceptibility is defined with respect to externally applied magnetic fields, but it is also of interest to consider the relationship between the induced magnetization of a body and its internal H field, which is the sum of the external field and the demagnetizing field due to 'poles' on the surface: $\vec{H}_i = \vec{H}_0 + \vec{H}_d$. The ratio M_i/H_i is the intrinsic susceptibility k_i. For most materials, k_i is essentially identical to k, but for those with high magnetization intensities, the demagnetizing fields become large enough that observed susceptibilities are self-limiting:

$$k_i = \frac{M_i}{H_i} = \frac{M_i}{H_0 + H_d} = \frac{M_i}{H_0 - NM_i} = \frac{kH_0}{H_0 - NkH_0} = \frac{k}{1 - Nk} \quad [6]$$

and thus,

$$k = \frac{k_i}{1 + Nk_i} \quad [7]$$

As $k_i \to 0$, $k \to k_i$, and as $k_i \to \infty$, $k \to 1/N$, which is the 'self-demagnetization limit.' Because N is anisotropic and related to particle shape, measured susceptibilities of strongly magnetic particles commonly exhibit a shape-related anisotropy.

Under certain conditions (including absence of free electrical currents), the flux density B is related to magnetic field H in a manner analogous to the relation of groundwater flux to hydraulic gradient and permeability (Darcy's law): $B = \mu H$, where the magnetic permeability μ is related to susceptibility and to the fundamental constant μ_0 (the permeability of free space). Within a material whose magnetization is proportional to the (internal) magnetic field ($M_i = k_i H_i$),

$$\begin{aligned}\vec{B}_i &= \mu_0\left(\vec{H}_i + \vec{M}_i\right) = \mu_0\left(\vec{H}_i + k_i \vec{H}_i\right) = \mu_0(1 + k_i)\,\vec{H}_i \\ &= \mu_0\mu_r\,\vec{H}_i = \mu\,\vec{H}_i\end{aligned} \quad [8]$$

The relative permeability μ_r, like volumetric susceptibility, is dimensionless; the absolute permeability μ has the same units as μ_0 (henry m^{-1}).

11.05.2.2 Classes of Magnetic Behavior

Magnetism results from the orbital motions of electrons around a nucleus and from the intrinsic property of electrons called spin. Each source is associated with a magnetic moment, and how the moments are combined to produce a net magnetic moment for an ion, molecule, or compound is governed by quantum mechanics. Atomic moments with orbital and spin contributions are quantized in multiples of the Bohr magneton (μ_B):

$$\mu_B = \frac{eh}{4\pi m_e} = 9.274 \times 10^{-24}\ \mathrm{A\,m^2} \quad [9]$$

where e and m_e are the charge and mass of an electron and h is Plank's constant. The spin and orbital components may combine in such a way to produce complete cancellation and a zero net atomic moment or only partial cancellation and a net atomic moment.

In addition, the nucleus also possesses a magnetic moment due to neutrons and protons, but nuclear magnetism is a much weaker effect than the atomic moments produced by electrons and does not contribute to the typical magnetic properties of materials of geophysical interest. However, the magnetic moment of protons is responsible for the phenomena of nuclear magnetic resonance (NMR), which is the working principle behind proton precession magnetometers (Turner et al., 2007) and surface and borehole NMR methods (Yaramanci and Müller-Petke, 2013).

Magnetism can be divided into five fundamental types based on the interactions of electrons with applied fields or with each other: (1) diamagnetism results from the orbital motions of electrons in the presence of an applied field; (2) paramagnetism results from the atomic moments associated with unpaired electrons in partially filled orbitals that can be aligned by an applied field but randomized by temperature; and (3–5) FM, antiferromagnetism (AF), and ferrimagnetism (FI) result from the spontaneous alignment of electron moments in the absence of an applied field.

Unlike diamagnetism and paramagnetism where spins do not interact with each other, materials exhibiting FM or AF have strong interactions between unpaired spins. This interaction produces a spontaneous ordering of electron moments in the absence of an applied field. The physical origin of the strong interactions between spins, called exchange interactions, comes from the quantum mechanical nature of electrons brought about from electron sharing and overlapping orbitals in crystals. Exchange interactions are a consequence of the Coulomb repulsion between two electrons and the Pauli exclusion principle, which forbids two electrons from occupying the same space with identical quantum states. The exchange forces are much stronger than the classical dipole–dipole interaction between two magnetic ions. Exchange interactions can be positive (FM) or negative (AF). They can also act directly between nearest-neighbor ionic moments (e.g., in metals) or indirectly through neighboring anions (e.g., O^{2-}, S^{-2}) in ionic compounds (e.g., oxides and sulfides) where the ionic moments are too far apart for direct interactions to be effective. The latter type of exchange interaction is called superexchange. In most naturally occurring magnetic oxides and sulfides, the iron ions carry the net spin magnetic moments associated with the number of unpaired electrons in the 3d orbitals of $Fe^{3+}(3d^5)$ and $Fe^{2+}(3d^6)$. The theory of magnetic exchange provides the basis for understanding magnetization and the different types of magnetic ordering in solids (**Figure 1**). Further information on magnetic ordering in solids and minerals can be found in Dunlop and Özdemir (1997).

11.05.2.2.1 Diamagnetism and paramagnetism

Diamagnetism is a property of all matter, related to the precession of electron orbits in an applied field; the precession is equivalent to a circulating electrical current, which produces a weak magnetic moment antiparallel to the field. The induced magnetization of bulk diamagnetic material is a linear function of applied field ($M_i = kH$), with k small, negative, and essentially independent of temperature and applied field strength. Diamagnetic susceptibility is much less variable than para- or ferromagnetic susceptibilities: for most materials, k_d is between about -10 and -20 microSI (χ_d between about -5×10^{-9} and $-8 \times 10^{-9}\ \mathrm{m^3\,kg^{-1}}$); bismuth is the strongest known diamagnet, with $k_d \sim -165$ microSI. Diamagnetism is for the

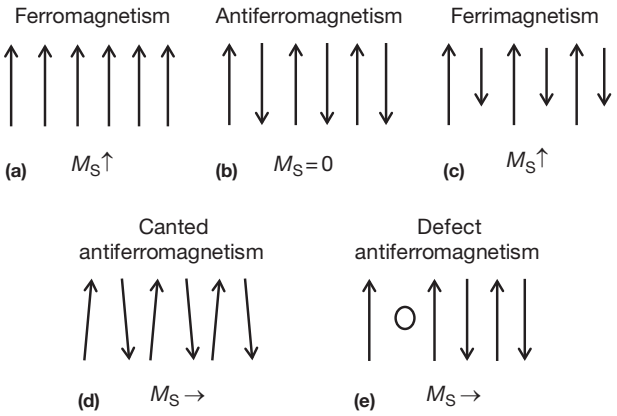


Figure 1 Types of magnetic ordering showing the alignment of exchange-coupled spin moments for (a) ferromagnetism, (b) antiferromagnetism, (c) ferrimagnetism, (d) canted antiferromagnetism, and (e) defect antiferromagnetism (circle represents a lattice defect or missing spin). The direction of the net magnetic moment, M_S, is indicated by the arrows.

most part isotropic, although calcite has a significant (~10%) variation in k_d according to crystallographic orientation (e.g., Schmidt et al., 2006).

Paramagnetism originates in atoms of certain transition metals (Fe, Ni, Mn, Co, and Cr) or rare earths (e.g., Gd, Dy, and Ho) that have uncompensated electron spins and therefore have permanent atomic magnetic moments. The atomic moments behave more or less independently in paramagnetic materials; their alignment with the ambient magnetic field is disrupted by random thermal perturbations, and paramagnetic susceptibility consequently varies inversely with absolute temperature according to the Curie law: $k_p = C/T$, where the Curie constant C depends on the concentration of magnetic ions in the bulk material and on their magnetic moments. When the magnetic ions are not completely independent but weakly coupled by a molecular field H_m whose intensity increases in proportion to the bulk magnetization of a paramagnetic material ($H_m = \lambda M$), the temperature dependence is modified, following the 'Curie–Weiss' law:

$$M = \frac{C(H_0 + H_m)}{T} = \frac{C(H_0 + \lambda M)}{T} = \frac{CH_0}{T - \lambda C} \qquad [10]$$

and

$$k = M/H_0 = \frac{C}{T - \lambda C} = \frac{C}{T - \theta} \qquad [11]$$

where θ is referred to as the 'paramagnetic Curie temperature.' It typically has values between −20 and +20 K (e.g., Rochette et al., 1992) for common Fe-bearing silicate minerals.

At sufficiently low temperatures (e.g., below 10 K) and in sufficiently strong applied fields (several Tesla), paramagnetic magnetization becomes a nonlinear function $M_p(H_0)$ as the alignment of atomic moments begins to approach saturation. This is described by the Langevin function (or the Brillouin function in quantum theory), which will be discussed in Section 11.05.4.4.

In geologic materials, paramagnetism derives mainly from Fe^{2+}, Fe^{3+}, and Mn^{2+}. To a very good approximation, paramagnetic susceptibility can be calculated (Syono, 1960) when the concentrations of these three ions are known:

$$k = \mu_0 \left(x_{Fe^{2+}} p^2_{Fe^{2+}} + x_{Fe^{3+}} p^2_{Fe^{3+}} + x_{Mn^{2+}} p^2_{Mn^{2+}} \right) / (3k_B T) \qquad [12]$$

where, for each cation, x is the number of ions per cubic meter and p is the atomic moment (~4 or 5 Bohr magnetons). Room-temperature susceptibilities of paramagnetic minerals commonly range up to $k_p \sim 1000$ microSI, or $\chi_p \sim 500 \times 10^{-9}\ m^3\ kg^{-1}$ (e.g., Borradaile and Jackson, 2010). Paramagnetic susceptibility can be strongly anisotropic, with directional variability commonly reaching 20–30% (e.g., Borradaile and Jackson, 2010; Rochette et al., 1992).

11.05.2.2.2 Ferromagnetism

Ferromagnetic materials have positive exchange interactions producing a parallel alignment of spins (**Figure 1(a)**) and a large net magnetization called the spontaneous or saturation magnetization (M_S). The magnitude of M_S at 0 K is dependent on the total spin magnetic moments of electrons in the FM material. For temperatures greater than 0 K, the degree of alignment, and hence $M_S(T)$, depends on temperature. Heating above a critical temperature, called the Curie temperature (T_C), causes the magnetic ordering to break down with a reversible loss of spontaneous magnetization. At temperatures above T_C, exchange interactions are no longer effective in keeping spins aligned against thermal perturbations, and each spin is free to respond individually to an external magnetic field. In this state, the material is paramagnetic following the Curie–Weiss law, exhibiting a weak magnetization in an applied field that vanishes when the field is removed. The Curie temperature is therefore a transition between ferromagnetic and paramagnetic behaviors. Most importantly, materials in the FM state have the ability to acquire remanence and be permanently magnetized. Curie temperatures vary over a wide temperature range from < 10 K up to 1388 K for cobalt. Only four elements (Fe, Ni, Co, and Gd) but numerous industrial metallic alloys are ferromagnetic at room temperature. Naturally occurring metallic iron and ferromagnetic metal alloys of Fe–Ni–Co–Cu are found in meteorites and lunar samples (Rochette et al., 2009) and in ultramafic rocks and serpentinites (Haggerty and Toft, 1985; Klein and Bach, 2009). Metallic iron and FeSi alloys are found in some fulgurites (Essene and Fisher, 1986).

11.05.2.2.3 Antiferromagnetism

Antiferromagnetic materials have negative exchange interactions producing an antiparallel alignment of spins (**Figure 1(b)**) with the atomic moments forming two equivalent but oppositely oriented magnetic sublattices (A and B). Each sublattice has a spontaneous magnetization oriented in opposite directions ($M_A = -M_B$) at all temperatures so that the sublattice configuration produces a zero net magnetization ($M_S = M_A + M_B = 0$). Hence, AF phases are incapable of acquiring remanence. Like FM, the magnetic order breaks down, and the sublattice spontaneous magnetization disappears at a critical temperature, which for antiferromagnets is the Néel temperature (T_N). Above T_N, paramagnetism is observed. Below T_N, even though the spin cancellation is perfect with $M_S = 0$ when $H_a = 0$, applied fields can slightly perturb the antiparallel

spin alignment and produce a weak induced magnetization and a positive susceptibility. The susceptibility is temperature-dependent and increases with T until it reaches a maximum at T_N, above which susceptibility then decreases according to the Curie–Weiss law of paramagnetism (eqn [11]). The resulting peak in susceptibility at the Néel temperature is a diagnostic feature of AF behavior. Many more compounds exhibit AF than FM, including naturally occurring minerals containing Fe and Mn in the form of oxides, hydroxides, sulfides, carbonates, and silicates. However, most of these minerals have Néel temperatures well below room temperature and are therefore paramagnetic at ambient temperatures. The notable exceptions are hematite with $T_N = 953$ K and goethite with $T_N = 393$ K, but these minerals have imperfect AF configurations as will be described later.

11.05.2.2.4 Ferrimagnetism

FI is a special case of AF (Figure 1(c)), with the net magnetization of the two sublattices antiparallel but unequal ($M_A \neq -M_B$). This produces a net spontaneous magnetization like in FM that disappears above a critical temperature also called the Curie temperature (T_C). In fact, many of the typical magnetic properties associated with FM, such as remanence and hysteresis, are also found in ferrimagnets. The spin imbalance in FI materials is due to different populations of diamagnetic (e.g., Mg^{2+}, Al^{3+}, and Ti^{4+}) and magnetic (Fe^{3+} and Fe^{2+}) cations occupying the A and B sites. Magnetite is the archetypal example of a ferrimagnetic material ($T_C = 853$ K), which for many years was assumed to be ferromagnetic until the theoretical work by Néel in the 1930s and 1940s provided the basis for understanding AF and FI. Almost all naturally occurring magnetic minerals are ferrimagnetic with the most common being the Fe–Ti oxides. Depending on composition and cation substitutions, most of these phases have Curie temperatures above room temperature yielding remanence and high susceptibilities at room temperature.

11.05.2.2.5 Canted and defect antiferromagnetism

Small deviations from perfect AF can occur for several reasons, all leading to a weak magnetization in zero field (Figure 1(d) and 1(e)) that is $<1\%$ of the spontaneous magnetization observed in typical FM and FI materials. Canted AF occurs when the spins of the two sublattices are not completely collinear but are slightly tilted or canted out of alignment ($<1^\circ$) producing a weak magnetization perpendicular to the AF axis. This type of magnetic ordering is a consequence of superexchange and certain types of crystal symmetries. Hematite is an example of a canted AF mineral and the weak magnetization disappears at its antiferromagnetic Néel temperature. In defect AF, the spin imbalance can be caused by crystal defects, impurity cations (e.g., Al^{3+}), poor crystallinity, or small particle sizes (<100 nm) and is likely responsible for the weak magnetization at room temperature observed in goethite. The defect AF moment also disappears at the AF Néel temperature.

11.05.2.3 Magnetic Units

Although the SI (Système Internationale) units of measure were adopted more than 50 years ago, the older cgs (centimeter–gram–second) system is still rather commonly used in magnetism, and converting units and equations from one system to the other is not as straightforward as one might expect, for various reasons (see Table 1 for conversion factors). For one, the cgs system itself was not coherent, having different subsystems optimized for working with particular sets of phenomena; these included the esu, emu, and Gaussian–cgs systems of units, the latter being the one that is still in common use. Physical constants and units were defined in convenient ways for each system: For example, μ_0 is dimensionless, with a value of unity in the emu and Gaussian–cgs systems, and ε_0 (permittivity of free space) was similarly defined in the esu system. The SI system provides a self-consistent set of units and constants for universal application, and therefore, the dimensional relationships between electromagnetic quantities are sometimes different in SI and in cgs. For example, B and H are dimensionally equal in the Gaussian–cgs system, but not in SI. Further, although each system is based on a small number of fundamental units, the SI (meter–kilogram–second–ampere) system includes electrical current in addition to length, mass, and time. Beyond that, in both systems, the commonly used units for magnetic moment, field, and flux density ignore or obscure the fundamental system units. For example, the cgs unit for B field is called the 'gauss,' rather than the equivalent expression in fundamental units, '$cm^{-1/2}\ g^{1/2}\ s^{-1}$,' and likewise in SI, it is called 'tesla' or

Table 1 Magnetic units and conversions

Parameter	*SI unit*	*CGS unit*	*Conversion*
Magnetic moment	$A\ m^2$	emu	$1\ Am^2 = 10^3$ emu
Magnetization			
By volume	$A\ m^{-1}$	emu cm^{-3}	$1\ A\ m^{-1} = 10^{-3}$ emu cm^{-3}
By mass	$A\ m^2\ kg^{-1}$	emu g^{-1}	$1\ A\ m^2\ kg^{-1} = 1$ emu g^{-1}
Magnetic field (H)	$A\ m^{-1}$	Oersted (Oe)	$1\ A\ m^{-1} = 4\pi \times 10^{-3}$ Oe
Magnetic induction (B)	Tesla (T)	Gauss (G)	$1\ T = 10^4$ G
Susceptibility	m^3	emu Oe^{-1}	$1\ m^3 = 10^6/4\pi$ emu Oe^{-1}
By volume ($k = M/H$)	–	emu $cm^{-3}\ Oe^{-1}$	1 SI $= 1/4\pi$ emu $cm^{-3}\ Oe^{-1}$
By mass ($\chi = M/H$)	$m^3\ kg^{-1}$	emu $g^{-1}\ Oe^{-1}$	$1\ m^3\ kg^{-1} = 10^3/4\pi$ emu $g^{-1}\ Oe^{-1}$
Permeability ($\mu = B/H$)	$H\ m^{-1}$	G Oe^{-1}	$1\ H\ m^{-1} = 10^7/4\pi$ G Oe^{-1}
Permeability of free space (μ_0)	$H\ m^{-1}$	1	$4\pi \times 10^{-7}\ Hm^{-1} = 1$ CGS

emu, electromagnetic unit; 1 tesla (T), kg $A^{-1}\ s^{-1}$; 1 henry (H), 1 newton A^{-2}.

'webers m^{-2}' or occasionally '$J\,A^{-1}\,m^{-2}$,' but never its fundamental equivalent '$kg\,A^{-1}\,s^{-2}$.' Finally, SI is a 'rationalized' unit system, in which the factor 4π, representing the flux through a spherical surface due to a central unit charge or pole, is included in the values of the constants rather than appearing explicitly in Maxwell's equations.

11.05.3 Magnetic Ordering and Magnetic Minerals

FM and FI are uncommon mineral properties, and of the thousands of known minerals, just a few dozen exhibit these properties. The number is reduced even further if we identify only those compositions that have $T_c > 273$ K and are therefore capable of producing remanent or induced magnetic anomalies. Magnetic minerals in rocks also typically occur in low modal abundance (< 5 vol.%).

Almost all natural magnetic minerals are Fe-bearing, belonging to the Fe–Ti oxides, Fe sulfides, Fe oxyhydroxides, and Fe–Ni metal groups. By far, the most common rock-forming magnetic minerals belong to the Fe–Ti oxides and their compositions can be summarized in $FeO–Fe_2O_3–TiO_2$ ternary diagram (Figure 2). Within the ternary diagram are the titanomagnetite ($Fe_{3-x}Ti_xO_4$, $0 \le x \le 1$), titanohematite ($Fe_{2-y}Ti_yO_4$, $0 \le y \le 1$), and titanomaghemite (oxidized titanomagnetites) series. Another series on the diagram is represented by the orthorhombic minerals in the pseudobrookite group; however, this group of minerals is paramagnetic at room temperature. Moreover, the majority of the natural magnetic minerals fall mainly in the quadrilateral defined by magnetite, ulvöspinel, pseudobrookite, and hematite.

The occurrence and abundance of magnetic minerals in rocks and unconsolidated materials are controlled by a variety of complex processes including magma composition, subsolidus and exsolution reactions during cooling of plutonic rocks, low-temperature oxidation of volcanic rocks, serpentinization of ultramafic rocks, and reduction diagenesis in marine and lake sediments. More detailed accounts of magnetic mineralogy can be found in O'Reilly (1984), Lindsley (1991), Dunlop and Özdemir (1997), and Harrison and Putnis (1999).

11.05.3.1 Iron Oxides

11.05.3.1.1 Magnetite

Magnetite (Fe_3O_4) is ubiquitous in the crust and sediments and can occur as either a primary or secondary mineral in igneous, metamorphic, and sedimentary rocks, as well as in sediments and soils. Nanoparticles (< 100 nm) of magnetite are also produced by a variety of microorganisms in sedimentary environments (Bazylinski and Moskowitz, 1997). Although magnetite is common, it is typically a minor accessory mineral, rarely exceeding 1% by volume. Nevertheless, magnetite has the largest saturation magnetization ($M_S = 480$ kA m^{-1} or 92 $Am^2\,kg^{-1}$) of all known natural magnetic minerals, so even trace amounts (< 0.1%) can be detected easily with laboratory magnetic measurements.

Magnetite is cubic with a spinel structure consisting of a close-packed arrangement of oxygen ions where small metal cations occupy interstitial sites. The interstices form two nonequivalent sites with tetrahedral (A sites) or octahedral (B sites) oxygen coordination; the two sites form two magnetic sublattices such that the spins in the A sublattice are coupled antiparallel to those in the B sublattice. However, the Fe ions are distributed unequally between the two sites with Fe^{3+} in the tetrahedral sites and $Fe^{2+}$and Fe^{3+} in the octahedral sites, yielding a structural formula for magnetite, $[Fe^{3+}]_A\,[Fe^{3+}, Fe^{2+}]_B O^{2-}{}_4$. Hence, magnetite is ferrimagnetic.

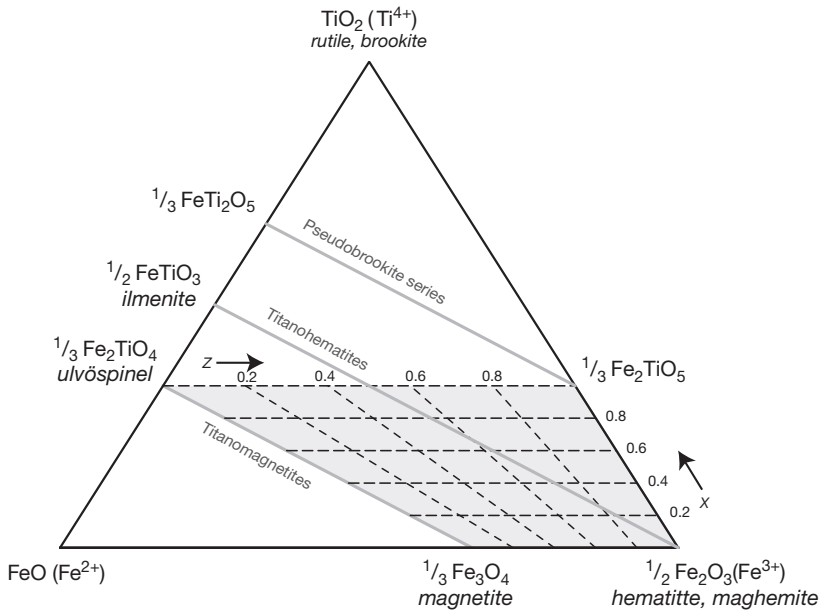


Figure 2 The $FeO–Fe_2O_3–TiO_2$ ternary diagram of rock magnetism. Important iron–titanium oxide magnetic minerals include the magnetite–ulvöspinel (titanomagnetite) and hematite–ilmenite (titanohematite) solid solution series; variably oxidized titanomagnetites (titanomaghemites) have compositions in the gray shaded area. Redrawn after O'Reilly W (1984) *Rock and Mineral Magnetism*. Glasgow: Blackie.

Magnetite exhibits both a structural phase transition at $T=125$ K, called the Verwey transition (T_v), and a magnetic isotropic point (T_i) at $T=130$ K (e.g., Dunlop and Özdemir, 1997). At the Verwey transition, stoichiometric magnetite undergoes a first-order crystallographic transition, in which the lattice structure changes from an inverse cubic spinel ($T>T_v$) to monoclinic ($T<T_v$) symmetry (e.g., Kasama et al., 2012; Piekarz et al., 2007; Walz, 2002). At the magnetic isotropic point, which occurs at a temperature slightly above T_v, the cubic magnetocrystalline anisotropy constant (K_1) becomes zero as it changes sign, and the easy directions of magnetization change their orientation from [1 1 1] above T_i to [1 0 0] below T_i (see **Section 11.05.4.3.1**). Both critical transitions produce a distinctive magnetic signature in remanence or magnetic susceptibility. For example, remanence given above or below T_v or T_i will be partially demagnetized upon cycling through these transitions. Likewise, magnetic susceptibility is a maximum at T_i, where the anisotropy energy is a minimum, and drops sharply below T_v as the anisotropy of monoclinic magnetite markedly increases. This behavior provides a diagnostic signature for stoichiometric magnetite. Slight deviations from stoichiometry (cation deficiency) or minor amounts of cation substitution (e.g., Ti^{4+} or Al^{3+}) can reduce the Verwey transition below 100 K or suppress it entirely (Moskowitz et al., 1998; Özdemir et al., 1993).

Minerals that crystallize in the spinel structure have the general formula AB_2O_4 where A and B can be divalent, trivalent, or quadrivalent cations in charge-balanced proportions. In 'normal' spinels (like the mineral spinel, $MgAl_2O_4$), the A cations reside in the tetrahedral A sites and the B cations are located in the octahedral B sites; in 'inverse' spinels (like magnetite), the A sites contain B cations and the A cations are located in the B sites. The most important solid solution series is that containing the titanomagnetites, which will be discussed in the next section. However, magnetite forms limited or complete solid solutions at elevated temperatures with several other spinel end-members containing Al, Mg, Mn, Cr, and Ni, such as hercynite ($FeAl_2O_4$), trevorite ($NiFe_2O_4$), magnesioferrite ($MgFe_2O_4$), jacobsite ($MnFe_2O_4$), and chromite ($FeCr_2O_4$). Usually, the magnetite-rich compositions are ferrimagnetic at ambient temperatures (i.e., $T_C>273$ K) with high saturation magnetization values but never larger than that of pure magnetite (**Table 2**). Although spinels are common in mafic and ultramafic rocks, Fe-rich spinels other than titanomagnetite almost never dominate the remanence. An exception may be in certain archaeological copper slag deposits where Fe–Mn oxides carry the remanent signal (Shaar et al., 2010).

Table 2 Magnetic properties of selected magnetic minerals

Mineral	*Composition*	T_C *(K),* T_n *(K)*	M_S *(A m^2 kg^{-1})*	*Type of magnetic ordering*
Oxides				
Magnetite	Fe_3O_4	853	92	FI
Maghemite	γ-Fe_2O_3	863–948	73	FI
Jacobsite	$MnFe_2O_4$	673	77	FI
Trevorite	$NiFe_2O_4$	858	51	FI
Magnesioferrite	$MgFe_2O_4$	713	27	FI
Titanomagnetite ($x=0.6$)	$Ti_{0.6}Fe_{2.4}O_4$	423	25	FI
Hematite	α-Fe_2O_3	948	0.4	Canted AF
Titanohematite ($y=0.7$)	$Ti_{0.7}Fe_{1.3}O_3$	~430	~30	FI
Ilmenite	$FeTiO_3$	40	PM	AF
Ulvöspinel	Fe_2TiO_4	120	PM	AF
Fe–chromite	$FeCr_2O_4$	88	PM	FI
Sulfides				
Pyrrhotite (monoclinic)	Fe_7S_8	593	20	FI
Pyrrhotite (hexagonal)	Fe_9S_{10}	538	0	FI
Troilite	FeS	578	0	AF
Greigite	Fe_3S_4	>623	59	FI
Daubreelite	$FeCr_2S_4$	~170	PM	FI
Oxyhydroxides				
Goethite	α-FeOOH	393	<0.5	Defect-AF
Lepidocrocite	γ-FeOOH	52	PM	Defect-AF
Feroxyhyte	δ-FeOOH	~450	~12	FI(?)
Ferrihydrite	$Fe_5HO_8 4H_2O$			
Metals, alloys, and other phases				
Iron	α-Fe	1043	218	FM
Nickel	Ni	627	55	FM
Awaruite	Ni_3Fe	960	120	FM
Wairauite	CoFe	1256	235	
Siderite	$FeCO_2$	35	PM	AF

FI, ferrimagnetism; FM, ferromagnetism; AF, antiferromagnetism; M_S, saturation magnetization at 300 K; T_C, Curie temperature; T_n, Néel temperature; PM, paramagnetic at room temperature but phase orders magnetically at low temperatures.

11.05.3.1.2 Titanomagnetite

Titanomagnetites are common magnetic minerals in igneous and metamorphic rocks. Titanomagnetites ($Fe_{3-x}Ti_xO_4$, $0 \le x \le 1$) form a complete solid solution above ~923 K between end-members magnetite (Fe_3O_4, $x=0$) and ulvöspinel (Fe_2TiO_4, $x=1$). The compositional parameter x gives the fraction of ulvöspinel ($(1-x)Fe_3O_4$ xFe_2TiO_4) or the mole fraction of Ti^{4+} in a particular composition. For example, mid-ocean ridge basalts (MORBs) have titanomagnetite compositions near $x \sim 0.6$. Like magnetite, titanomagnetites are cubic spinels where the substitution of Ti for Fe occurs by replacing $2Fe^{3+}$ for $Ti^{4+}+Fe^{2+}$. Many intrinsic magnetic and physical properties of titanomagnetites, including saturation magnetization, Curie temperature, magnetocrystalline anisotropy, magnetostriction, and electrical conductivity depend on a variety of interrelated crystal chemical properties. These include chemical composition, cation ordering, and defect content, each of which may be modified during initial petrogenesis or subsequent thermal history. The effect of Ti substitution causes both the M_S and T_C to decrease approximately linearly with the x-parameter (e.g., Lattard et al., 2006). Titanomagnetites with x greater than approximately 0.75 have Curie temperatures below 273 K and hence are paramagnetic at ambient temperatures (Figure 3). Cations other than Ti can also be substituted into the spinel crystal structure and natural titanomagnetites also commonly contain minor amounts of Al^{3+}, Mg^{2+}, and Mn^{2+}. The main effects of these additional cations are to lower both T_C and M_S from the value in the ideal stoichiometric titanomagnetite.

11.05.3.1.3 Maghemite and titanomaghemite

Fine-grained (titano)magnetite is subject to low-temperature oxidation ($T < 473$ K), a process caused by the diffusion of Fe^{2+} ions toward the surface of the crystal, where they become oxidized (Dunlop and Özdemir, 1997). Low-temperature oxidation of (titano)magnetite to a single-phase oxidation product is called maghemitization or titanomaghemitization. The oxidation products are referred to as nonstoichiometric or cation-deficient spinels (titanomaghemites) meaning that the metal/oxygen ratio is less than that (3:4) for ideal stoichiometry. There is also no change in crystal structure during oxidation. Maghemite (γ-Fe_2O_3) is the fully oxidized version of magnetite. It has the same inverse spinel crystal structure as magnetite but contains no divalent iron and therefore has the same chemical composition as hematite (α-Fe_2O_3). To maintain charge balance, the oxidation of Fe^{2+} to Fe^{3+} is accompanied by the formation of cation vacancies on the B sites such that 1/6 of the Fe sites are vacant. Maghemite is ferrimagnetic like magnetite with a saturation magnetization at room temperature of approximately 70 $A\,m^2\,kg^{-1}$ (~340 $kA\,m^{-1}$). However, maghemite is thermodynamically unstable and will transform to hematite when heated to $T > 573$ K; the inversion temperature is controlled in part by particle size and by Al substitution (de Boer and Dekkers, 1996). This makes the determination of the Curie temperature problematic but estimates indicate it is between 863 and 949 K (de Boer and Dekkers, 1996; Gehring et al., 2009; Özdemir and Banerjee, 1984). Maghemite has no low-temperature ($T < 300$ K) remanence transitions.

Titanomagnetite will undergo low-temperature oxidation to produce single-phase titanomaghemites. However, in this case, the term titanomaghemite refers to compositions with a range of oxidation states from near stoichiometric to fully oxidized. The degree of oxidation is described by the z-parameter ($0 \le z \le 1$), which is the fraction of Fe^{2+} in the original titanomagnetite that has been converted to Fe^{3+}, with $z=1$ being the fully oxidized composition (e.g., O'Reilly, 1984). Titanomagnetites in MORB readily undergo low-temperature oxidation to titanomaghemite, and this is a common form of alteration at ambient temperatures. At high temperatures ($T > 1200$ K), titanomagnetites can accommodate a limited degree of nonstoichiometry of a few percent ($z < 0.1$) within their stability field before inverting to a multiphase assemblage (Aragon and McCallister, 1982; O'Reilly, 1984). These phases are termed nonstoichiometric titanomagnetite instead of titanomaghemite.

Oxidation of titanomagnetites to titanomaghemites is accompanied by a rise in Curie temperature and a decrease in saturation magnetization. Like maghemite, titanomaghemites are thermodynamically unstable and heating above about 625 K produces a variety of inversion products depending on the composition of the spinel phase prior to inversion. Generally, inversion produces an Fe-rich spinel phase (ferrimagnetic at room temperature) and a Ti-rich ilmenite phase (paramagnetic at room temperature). Hematite, TiO_2 (rutile or anatase), and pseudobrookite are inversion products of more highly oxidized compositions.

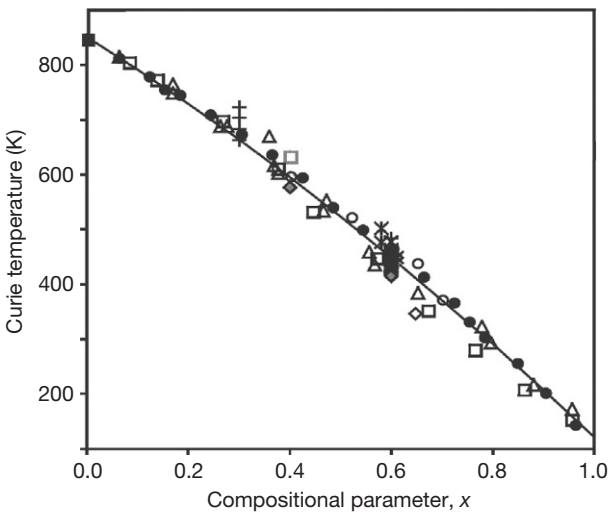


Figure 3 Curie temperatures in the magnetite–ulvöspinel (titanomagnetite) solid solution series for synthetic samples from various authors (see Lattard et al. (2006) for details). The solid curve is a best-fit polynomial: $T_C = -150x^2 - 580x + 851$. Figure modified from Lattard D, Engelmann R, Kontny A, and Sauerzapf U (2006) Curie temperatures of synthetic titanomagnetites in the Fe–Ti–O system: Effects of composition, crystal chemistry, and thermomagnetic methods. *Journal of Geophysical Research-Solid Earth* 111: B12S28. http://dx.doi.org/10.1029/2006JB004591.

11.05.3.1.4 Hematite

Hematite (α-Fe_2O_3) crystallizes in the corundum structure with oxygen ions in a hexagonal close-packed framework. The magnetic moments of the Fe^{3+} ions are ferromagnetically coupled within the c-planes, but antiferromagnetically coupled between the planes. However, due to spin canting

(Figure 1(d)), there is a weak FM ($M_S = 2$ kA m^{-1} or 0.38 m A m^2 kg^{-1}) observed above 262 K, which is ~0.4% that of magnetite. The Curie temperature is 953 K, the highest for all known Fe oxides. Below 262 K, the direction of the AF changes and becomes parallel to the *c*-axis; there is no spin canting and hematite becomes a perfect antiferromagnet. The spin–flop transition is called the Morin transition (T_M), although the exact temperature of the Morin transition is dependent on a number of variables including grain size, cation substitution, and lattice defects (Özdemir and Dunlop, 2006). As little as 1 mol% Ti substitution will reduce T_M below 10 K, while the transition completely disappears in nanosized particles of hematite (Morin, 1950; Nininger and Schroeer, 1978).

Hematite is the stable end product of oxidation of magnetite or the inversion of maghemite. It can also form during oxyexsolution of titanomagnetite during slow cooling or by inversion of titanomaghemite during heating. Weathering and dehydration reactions of iron oxyhydroxides, such as goethite and ferrihydrite, will also produce hematite. In soils, hematite (and goethite) can have concentrations of up to 5% by mass (Maher, 2011).

11.05.3.1.5 Titanohematite

Titanohematites ($Fe_{2-y}Ti_yO_3$, $0 \le y \le 1$) are rhombohedral minerals that form complete solid solution above ~1073 K (e.g., Harrison et al., 2007; McEnroe et al., 2009) between end-members hematite (α-Fe_2O_3, $y = 0$) and ilmenite ($FeTiO_3$, $y = 1$). The compositional parameter y gives the fraction of ilmenite ($(1-y)Fe_2O_3$ $yFeTiO_3$) or the mole fraction of Ti^{4+} in a particular composition. Hemoilmenite is another name used for this solid solution series, but the term also refers to exsolved grains containing fine-grained intergrowths of Fe-rich titanohematite and Ti-rich titanohematite (ferrian ilmenite) lamella. The hematite–ilmenite solid solution series shows a complex interplay of magnetic and crystal chemical features combining magnetic ordering, cation ordering, and subsolvus exsolution phenomena (e.g., Burton, 1991; Charilaou et al., 2012; Harrison et al., 2007). Compositions close to hematite ($y < 0.4$) have disordered cation distributions, canted antiferromagnetic spin structures, and weak magnetic moments similar to hematite. Titanohematites with compositions $y > 0.5$ have partially ordered cation distributions and ferrimagnetic spin structures with magnetizations reaching as much as 50% of the value for magnetite at room temperature. The Curie temperature decreases approximately linearly with composition (Figure 4), falling below room temperature at about $y = 0.8$. End-member ilmenite has a Néel temperature of 57 K.

The best-known, and perhaps, most infamous feature of this solid solution series is the property of self-reversed thermoremanent magnetization (TRM) (i.e., they acquire a remanent magnetization on cooling that is antiparallel to the applied field direction) observed in single-phase compositions containing 50–70% ($y = 0.5$–0.7) ilmenite (Bina et al., 1999; Hoffman, 1992; Ishikawa and Syono, 1963; Lagroix et al., 2004; Nord and Lawson, 1989, 1992; Prévot et al., 2001). Natural intermediate titanohematites are found in rapidly cooled dacitic pyroclastic rocks. More recent work on multiphase hemoilmenites has concentrated on their contributions to aeromagnetic anomalies from slowly cooled igneous and metamorphic rocks from the Adirondacks, Norway, and Sweden (Brown and McEnroe, 2008; Fabian et al., 2008; McEnroe and Brown, 2000; McEnroe et al., 2001a,b; Figure 5). These authors have shown that exsolved hemoilmenites can carry an intense stable natural remanence with high coercivity and are likely the sources for remanence-dominated negative anomalies in anorthosites and some crustal amphibolite and granulite terranes where these phases are present. As an historical note, the work by Balsley and Buddington (1958) in the Adirondack Mts. found a clear correlation between negative aeromagnetic anomalies and rocks with exsolved hemoilmenite compositions; this discovery featured prominently in the early debate over field reversals versus self-reversals in the late 1950s.

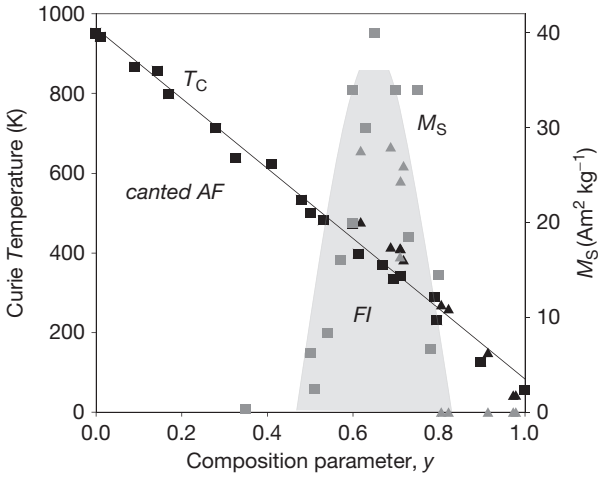


Figure 4 Curie temperature and saturation magnetization (at 300 K) in the hematite–ilmenite (titanohematite) solid solution series ($Fe_{2-y}Ti_yO_3$). Data from Hunt et al. (1995) denoted by squares and Burton et al. (2008) denoted by triangles. Black symbols = Curie temperature, gray symbols = saturation magnetization, and FI is the ferrimagnetic region at 300 K. The solid curve is a best-fit line for Curie temperature data: $T_C = -963.9 - 880.7y$. Figure modified from Hunt CP, Moskowitz BM, and Banerjee SK (1995) Magnetic properties of rocks and minerals. In: Ahrens TJ (ed.) *A Handbook of Physical Constants*, vol. 3. Washington, DC: American Geophysical Union, pp. 189–204.

The first-order observation is that the high natural remanent magnetizations (NRMs) in hemoilmenite are inconsistent with the bulk canted antiferromagnetic structure expected for the Ti-poor exsolution lamellae (McEnroe et al., 2001a). Instead, the explanation lies in an entirely new class of magnetic nanostructures and magnetic ordering, called lamellar magnetism (Harrison, 2006; McEnroe et al., 2009; Robinson et al., 2002, 2004, 2012). It is based on atomic and magnetic interactions at the nanoscale interfaces (10–20 nm) between Ti-poor titanohematite (canted antiferromagnet) and Ti-rich ilmenite (paramagnetic) lamellae. Combining electron microscopy observations, neutron diffraction, and Monte Carlo modeling, a ferrimagnetically Fe^{3+}–Fe^{2+}-ordered interface layer is predicted to form between the exsolution lamellae, which can give rise to the high magnetization of bulk samples (e.g., Harrison et al., 2007; Robinson et al., 2012). It is estimated that lamellar magnetism can result in saturation magnetization values of ~55 kA m^{-1} (22 times larger than hematite and 11% of the moment in magnetite) and remanent magnetization up to 30 A m^{-1} (Robinson et al., 2002).

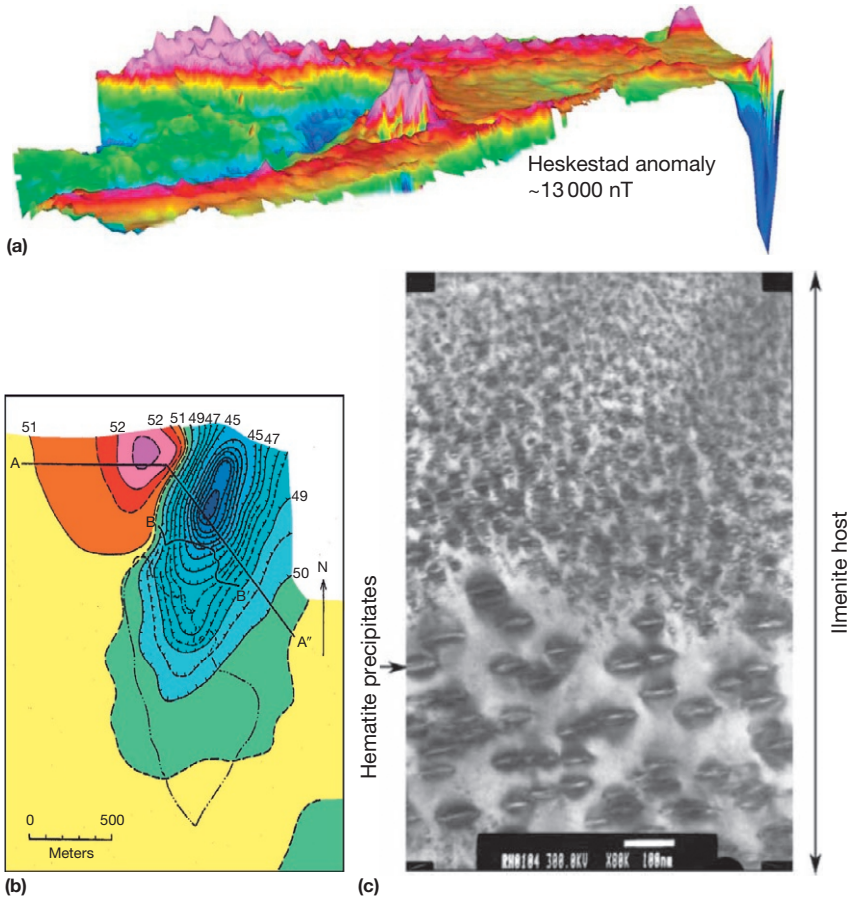


Figure 5 Aeromagnetic anomalies caused by nanoscale lamellae structures in hemoilmenite intergrowths. (a) Low-angle 3D image of Heskestad magnetic anomaly (southern Norway) from nearly 1 billion years old rocks. The negative anomaly is ~13 000 nT below background and is associated with an hemoilmenite norite unit of the Bjerkreim–Sokndal layered intrusion. (b) Magnetic anomaly map from a helicopter survey covering the same geographic area in (a). Solid contours are at 1000 nT intervals with additional dashed contours at 500 nT intervals. (c) Bright-field TEM image of the ilmenite host containing fine-scale hematite precipitates from the norite unit. Scale bar = 100 nm. Nanoscale compositional heterogeneities have been shown to produce high magnetizations. Figure modified from McEnroe SA, Harrison RJ, Robinson P, and Langenhorst F (2002) Nanoscale hematite–ilmenite lamellae in massive ilmenite rock: An example of 'lamellar magnetism' with implications for planetary magnetic anomalies. *Geophysical Journal International* 151: 890–912; McEnroe SA, Brown LL, and Robinson P (2004) Earth analog for Martian magnetic anomalies: Remanence properties of hemo-ilmenite norites in the Bjerkreim–Sokndal intrusion, Rogaland, Norway. *Journal of Applied Geophysics* 56 (3): 195–212.

11.05.3.2 Iron Sulfides and Other Magnetic Minerals

There are two iron sulfides, greigite (Fe_3S_4) and monoclinic pyrrhotite (Fe_7S_8), that are ferrimagnetic at room temperature and are geophysically important magnetic minerals like magnetite, titanomagnetite, and titanomaghemite. Pyrrhotite has the general formula of $Fe_{1-x}S$ ($0 \leq x \leq 0.13$) and crystallizes with the hexagonal NiAs structure consisting of alternating c-layers of Fe and S. The stoichiometric end-member ($x = 0$) is troilite (FeS), but all of the other variants ($x > 0$) are cation-deficient phases varying in composition between $Fe_{11}S_{12}$ and Fe_7S_8. The magnetic spin ordering in the series is similar to that of hematite but without the spin canting. Spins within the Fe planes are parallel but adjacent planes are antiferromagnetically coupled with antiparallel moments. Depending on the distribution of cation vacancies, the magnetic ordering can be antiferromagnetic (random distribution of vacancies) or ferrimagnetic (ordered distribution of vacancies). Troilite is antiferromagnetic at room temperature with a Néel temperature of 578 K. Ferrimagnetism in the series arises from the ordering of cation vacancies on alternating *c*-planes, which causes the magnetic sublattices to have unequal moments. Monoclinic pyrrhotite (F_7S_8) is the only form that is ferrimagnetic at room temperature ($M_S = 80$ kA m^{-1}, $T_C = 593$ K) and is the remanence-carrying phase in rocks and sediments. Hexagonal pyrrhotite (Fe_9S_{10} and $Fe_{10}S_{11}$) is antiferromagnetic at room temperature but exhibits FI in a narrow temperature window between about 473 and 543 K. Depending on composition, the hexagonal forms undergo a structural transition, known as the lambda transition around 473 K, where vacancy ordering occurs and FI develops until the Curie temperature is reached at $T \sim 543$ K (e.g., Dunlop and Özdemir, 1997).

Monoclinic pyrrhotite has a magnetic transition at $T = 34$ K (Besnus and Meyer, 1964; Rochette et al., 1990), now known as the Besnus transition, similar in style to the transitions observed in magnetite (Verwey transition) and hematite (Morin transition). Although the underlying cause for the

pyrrhotite transition is still debated (Kind et al., 2013), recent results suggest it is more like a structural phase transition and change in crystal symmetry similar to the Verwey transition than a spin–flop transition similar to the Morin transition (Wolfers et al., 2011). In any case, low-temperature cycling (5–300 K) of remanence is highly diagnostic of the presence of monoclinic pyrrhotite.

Naturally occurring pyrrhotite is usually a mixture of monoclinic and hexagonal varieties. Pyrrhotite has been found in trace amounts in igneous, metamorphic, and sedimentary rocks. It is common in low-grade metamorphic rocks, such as metasedimentary rocks associated with magmatic sulfide deposits, and is also a secondary phase in marine sediments. Under oxidizing conditions, pyrrhotite breaks down to various iron oxides and paramagnetic pyrite. Troilite is generally found in meteorites and lunar samples.

Greigite (Fe_3S_4) is an inverse spinel mineral and is the sulfide equivalent of magnetite (see Roberts et al. (2011) for a current review of the magnetism of greigite). It is ferrimagnetically ordered and strongly magnetic with $M_S = 59$ A m^2 kg^{-1} (239 kA m^{-1}), which is lower than both magnetite and maghemite. There is uncertainty in the Curie temperature because of the thermal breakdown of greigite upon heating to ~573 K, but current estimates indicate that it exceeds 623 K (Chang et al., 2008) and is likely higher than T_C for pyrrhotite or troilite. Greigite is a diagenetic mineral found in reducing (anaerobic) sedimentary environments where organic matter has been consumed by sulfate-reducing bacteria in the absence of oxygen. It is also produced by certain types of magnetotactic bacteria as intercellular nanocrystals (e.g., Bazylinski and Moskowitz, 1997). It is now recognized as an important remanence-carrying phase in the sedimentary record (Roberts et al., 2011). Unlike pyrrhotite and magnetite, greigite has no low-temperature (<300 K) remanence transitions (Roberts et al., 2011).

Goethite (α-FeOOH) is an iron oxyhydroxide mineral that crystallizes in the orthorhombic system. The magnetic structure is antiferromagnetic with the spins aligned parallel to the *c*-axis but there is a weak FM believed to be caused by random impurity cations or lattice defects that cause a slight spin imbalance between the alternating *c*-planes (**Figure 1(e)**). The Curie temperature is 393 K with saturation magnetization on the same order of magnitude as hematite (Dunlop and Özdemir, 1997; Özdemir and Dunlop, 1996). Besides the weak magnetization of goethite, the bulk AF ordering imparts an extremely high anisotropy resulting in nonsaturating isothermal remanence in DC fields in excess of 10 T (Rochette et al., 2005). The latter property sets it apart from hematite, which usually approaches saturation in DC fields below 2–3 T (Maher et al., 2004). Goethite also displays a small positive susceptibility typical of AFM materials and 2–3 orders of magnitude less than ferrimagnetic susceptibilities.

Goethite is a weathering product found in soils and sediments. It commonly forms as an ambient-temperature alteration product of iron-bearing chain and sheet silicate minerals (feldspar, pyroxene, amphibole, and mica) on the Earth's surface. Having once been formed, goethite can alter to hematite (α-Fe_2O_3) by dehydroxylation both at the Earth's surface and at moderate depth (as red sandstone and siltstones).

11.05.4 Induced and Remanent Magnetization

11.05.4.1 Magnetic Hysteresis

An important characteristic of FM and FI materials is the magnetization curve, $M(H)$, which relates the response of the induced magnetization to changing applied fields at constant temperature. The $M(H)$ curve is nonlinear, is usually anisotropic in single particles or crystals, and can be nonzero for $H = 0$. In high fields, it is reversible, but in low fields, the $M(H)$ curve shows hysteresis where the magnetization at a particular value of H depends on the previous magnetic field history. In other words, when the field is cycled between the limits $\pm H$, the $M(H)$ curves form a closed M–H loop, called a hysteresis loop (**Figure 6**). Several important magnetic parameters can be determined from the loop when H is large enough to insure saturation of the induced magnetization. The saturation magnetization (M_S) is the maximum induced magnetic moment that can be obtained in some large field (1–2 T); beyond this field value, no further increase in magnetization occurs. Upon reducing the field to zero, the magnetization does not go to zero but persists as a saturation remanence (M_{RS}). As the field is increased in the negative direction, the induced magnetization becomes zero at a field called the coercivity or coercive force (H_C). A related parameter is the coercivity of remanence (H_{CR}), which is the reverse field needed to reduce the saturation remanence to zero. It is always larger than the coercive force. Increasing the field further in the negative direction causes saturation in the negative direction, and cycling the field back to positive saturation traces out the major hysteresis loop. The hysteresis loop combines information on an intrinsic magnetic property, M_S, which depends on composition, and several extrinsic properties (M_R, H_C, H_{CR}, and χ_0), which depend on particle size, shape, geometric arrangement of particles, interparticle interactions, microstructure, and crystal defects. All hysteresis properties are temperature-dependent

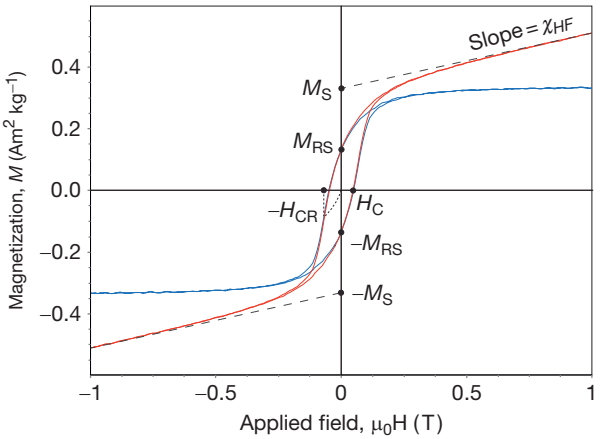


Figure 6 Major hysteresis loop, with contributions from ferri-, para-, and diamagnetic minerals, and calculated parameters as defined in the text (M_{RS}, saturation remanent magnetization; M_s, saturation magnetization; H_c, coercivity; H_{cr}, coercivity of remanence, χ_{HF} = high-field susceptibility). The high-field slope (in fields where the ferrimagnets are saturated) isolates the dia- and paramagnetic magnetization (represented by dashed line), which is subtracted from the measured loop (red) to obtain the ferrimagnetic loop (blue).

and approach zero as the temperature approaches T_C. Concise reviews of hysteresis measurements in rock magnetism are given by Krása and Fabian (2007), Jackson and Sølheid (2010), and Heslop and Roberts (2012).

11.05.4.2 Magnetic Domains

The theory of magnetic exchange explains the microscopic magnetic ordering in solids and properties such as Curie and Néel temperatures and spontaneous magnetization. However, it does not explain macroscopic or structure-sensitive properties such as magnetic remanence, coercivity, or the shape of a magnetic hysteresis loop. For example, the strength of exchange interactions between spins is equivalent to a field on the order of 1000 T, much larger than any field that can be produced in the laboratory and approximately 100 million times the strength of the geomagnetic field. Yet, remarkably, magnetization can be easily changed by the application of magnetic fields much less than 1 T as seen in hysteresis loops and even by the geomagnetic field. Furthermore, low DC or AC fields, or temperature, can bring the material into the demagnetized state where $M=0$ when $H=0$. To explain $M(H)$ behavior and still be consistent with exchange ordering below T_C, it is necessary to recognize that magnetic materials contain numerous internal micromagnetic structures called magnetic domains (e.g., Kittel, 1949). Each domain is spontaneously magnetized under the influence of exchange interactions to the saturation value, $M_S(T)$, but the magnetization vector varies from one domain to another. Changes in magnetization in relatively low fields are accommodated by the rearrangement of domains. Depending on the configuration of domains and the orientations of their magnetization directions, the average magnetization can be zero (demagnetized state), small with $M \ll M_S$, when $H=0$ (remanent state), or equal to M_S, when the field is large enough to align the domain magnetizations. Examples of idealized domain configurations are shown in Figure 7.

Magnetic domains are small (1–100 s μm), but much larger than atomic distances. The thin boundary layer between domains is called a domain wall where the magnetization gradually changes its orientation from one domain to the other. The width of domain walls is usually much less than domain widths, typically less than 500 nm depending on the magnetic exchange and anisotropy energies of the material. Domains and walls can be imaged directly using various magnetic imaging techniques (Figure 8). It is important to emphasize that the macroscopic breakup into domains should not be confused with the atomic-scale exchange ordering that produces parallel or antiparallel alignment of spins (FM, FI, or AF) that occurs within each domain. This distinction has sometimes been blurred in textbooks on applied geophysics.

Magnetic processes associated with remanence, coercivity, and hysteresis are strongly affected by the number of domains within a grain and by the response of domains to fields and temperature owing to the motion of domain walls and the rotation of domain moments. The micromagnetic states of ferromagnetic and ferrimagnetic particles are strongly dependent on particle size and can be divided into four classes:

Superparamagnetic (SPM) particles. Uniformly magnetized particles with particle volumes so small that the magnetization is unstable and spontaneously reverses its direction with relaxation times $\ll$ laboratory measurement times (<100 s). SPM grains have zero remanence and coercivity but enhanced susceptibility. Superparamagnetism occurs in nanosized ($<\sim$30 nm) particles for most magnetic minerals at room temperature.

Single-domain (SD) particles. Uniformly magnetized grains with relaxation times up to 1 Ga or more. SD grains can acquire an intense remanence in the geomagnetic field with large coercivities ($\gg$Earth's field) and retain stable long-term memory over geologic time. SD particles have the largest remanence per unit volume.

Multidomain (MD) particles. Nonuniformly magnetized particles (usually >1 μm) with weak remanence and low coercivities that make them inefficient recorders of the geomagnetic field. MD remanence can easily be remagnetized in nature but can have large magnetic susceptibilities.

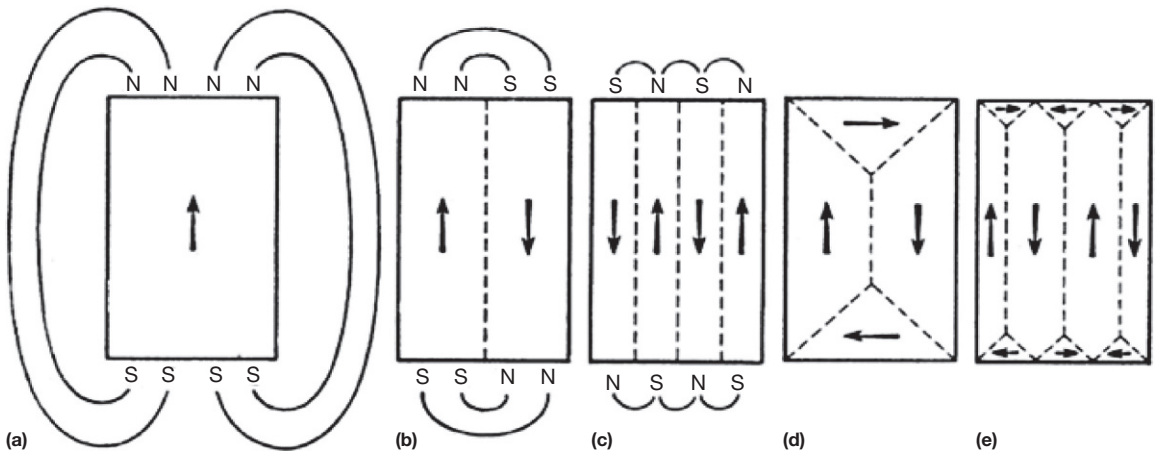


Figure 7 Idealized domain states representing (a) single-domain and (b–e) multidomain configurations. The subdivision into magnetic domains with antiparallel magnetizations decreases magnetostatic energy and reduces flux lines, allowing closure through neighboring domains (b–e). The MD configurations in (d) and (e) are two and four domain states containing closure domains, which close the flux entirely within the grain and reduce the magnetostatic energy to zero. The closure domains tend to be smaller and perpendicular to the main domains. Figure modified from Kittel C (1949) Physical theory of ferromagnetic domains, *Reviews of Modern Physics* 21: 541–583.

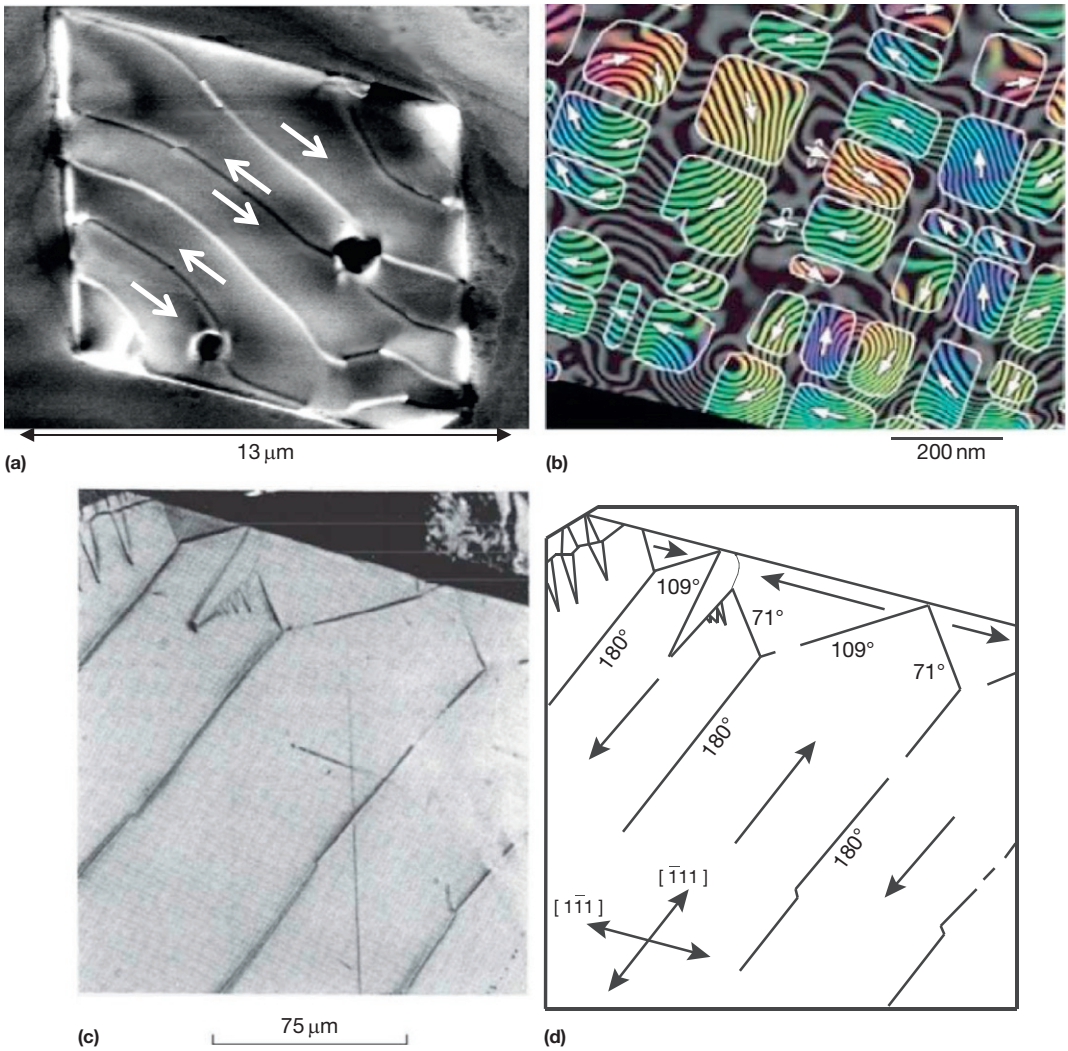


Figure 8 Magnetic microstructures in magnetite imaged with various techniques. (a) Magnetic force microscope image of magnetite grain containing seven domains. Black (white) contrast corresponds to plus (minus) magnetic surface poles and arrows represent direction of domain magnetization (Figure modified from Pokhil TG and Moskowitz BM (1997) Magnetic domains and domain walls in pseudo-single-domain magnetite studied with magnetic force microscopy. *Journal of Geophysical Research B: Solid Earth* 102(10): 22681–22694). (b) Electron holography image of magnetic microstructure in an intergrowth of magnetite (outlined in white) and nonmagnetic ulvöspinel showing magnetic interactions between magnetite lamellae. Colored contours and arrows indicate the distribution of magnetic flux within the magnetite lamellae. (Figure modified from Harrison RJ, et al. (2002).) (c) Bitter pattern image of natural single-crystal magnetite on (1 1 0) surface showing body domains and closure domains magnetized along (1 1 1) easy directions; (d) schematic image of domains in (c) showing direction of magnetization, domain wall types, and (1 1 1) easy directions. Figure modified from Özdemir Ö, Xu S, and Dunlop DJ (1995) Closure domains in magnetite, *Journal of Geophysical Research B* 100: 2193–2209.

Pseudo-single-domain (PSD) particles. Small MD particles that contain just a few domains or other magnetization structures (e.g., vortices) but exhibit SD-like properties of high remanence and long relaxation times. PSD grains can have stable remanence over geologic times and occur in particle size ranges that are most commonly observed in natural samples (~0.1–10 μm; the range varies according to mineralogy).

In a uniformly magnetized grain, exchange energy is minimized. However, to understand domain formation and superparamagnetism, it is necessary to introduce two additional energy terms, anisotropy energy and magnetostatic energy, which contribute to the total free energy of ferromagnetic particles.

11.05.4.3 Magnetism in Mineral Grains

11.05.4.3.1 Anisotropy energy

The energy of FM, FI, and AF materials depends on the direction of magnetization and this angular dependence is described by anisotropy energy. Anisotropy results in certain preferred directions with respect to the crystal lattice (magnetocrystalline anisotropy), the shape of the grain (shape anisotropy), or a stress direction (magnetoelastic anisotropy) along which M_S prefers to align. The preferred direction in which the energy is a minimum is called the easy direction. In contrast, the hard direction is an energy maximum and is avoided unless a high enough field is applied to force the magnetization to lie in a hard direction.

Magnetocrystalline anisotropy arises from the coupling between the spin moments and the crystalline lattice (spin–orbit coupling), and the easy and hard directions are related to the underlying crystal symmetry of the material. In cubic crystals, the magnetocrystalline anisotropy energy is given by

$$E_k = K_1 V\left(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_1^2\alpha_3^2\right) + K_2 V\left(\alpha_1^2\alpha_2^2\alpha_3^2\right) \qquad [13]$$

where K_1 and K_2 are the magnetocrystalline anisotropy constants determined from experimental measurements (e.g., M–H curves) and can be positive or negative, α_i's are the direction cosines of the magnetization relative to the $\langle 1\,0\,0\rangle$ crystal direction, and V is volume. Different magnetic materials will have different easy and hard directions depending on the sign of K_1 (assuming $K_2 = 0$). In materials with $K_1 < 0$ (e.g., magnetite above 130 K), the $\langle 1\,1\,1\rangle$ directions are the easy directions and the $\langle 1\,0\,0\rangle$ directions are the hard directions. In materials with $K_1 > 0$ (like Fe and Ti-rich titanomagnetites), $\langle 1\,0\,0\rangle$ are the easy directions and $\langle 1\,1\,1\rangle$ are the hard directions.

Another effect due to spin–orbit coupling is magnetostriction, which is the spontaneous elastic strain of the lattice that occurs upon magnetic ordering below T_C. A related phenomenon occurs when an external stress produces a stress-induced anisotropy that can rotate the magnetization away from the easy direction given by crystalline anisotropy. Therefore, stress can produce an easy axis of magnetization and the form of the anisotropy usually has a uniaxial symmetry. Depending on the material (via the magnitude and sign of the magnetostriction constants) and the stress direction, external stress can rotate M_S either toward or away from the compressive stress axis. A related phenomenon is piezomagnetism, which is the change in susceptibility or remanent magnetization when a magnetic material is subject to stress in the presence of a magnetic field (e.g., Gilder, 2007). The magnetocrystalline anisotropy and magnetostriction constants depend on mineral composition, crystal structure, temperature, and pressure but are independent of particle size. Values for these constants for common magnetic minerals can be found in Hunt et al. (1995) and Dunlop and Özdemir (1997).

Regardless of the origin of the anisotropy, to reverse the magnetization by an applied field from one easy axis to another requires rotation through a hard axis. Magnetic anisotropy creates an energy barrier that pins M along the easy axis until a large enough field is applied to cause an irreversible jump of M over the anisotropy barrier. Therefore, magnetic anisotropy strongly affects the shape of hysteresis loops, controls the coercivity, and is a measure of the stability of remanence against remagnetization. Without anisotropy, remanence would not be possible. Anisotropy is also of considerable practical importance because it is exploited in the design of many magnetic materials of commercial importance.

11.05.4.3.2 Magnetostatic energy

Magnetostatic energy arises when the magnetization of a body produces an external field. It is the sum of two terms, respectively, related to interaction of the magnetization with applied fields (Zeeman energy, eqn [3]) and internal fields (demagnetizing energy, eqn [14]). When the external field is zero, the internal field or the demagnetizing field remains. This contribution to the magnetostatic energy originates from classical dipole–dipole interactions where each dipole 'feels' the field created by all of the other dipoles. The magnetostatic energy is long-ranged, meaning that at any particular point, it depends on the magnetization distribution at every other point. The demagnetizing field arises whenever the magnetization has a component normal to a crystal surface ($\hat{n}\cdot\vec{M}\neq 0$) or when the magnetization is nonuniform such that $\nabla\cdot\vec{M}\neq 0$.

In a uniformly magnetized grain (Figure 7), uncompensated magnetic poles will form on a free surface and become a potent source of magnetostatic energy. An ellipsoid with volume V uniformly magnetized with magnetization M in zero applied field has a magnetostatic energy given by

$$E_D = {}^1\!/_2\mu_0 VNM^2 \qquad [14]$$

where N is the demagnetizing factor along the direction of M. The demagnetizing factor is determined by the shape of the sample. For a sphere $N = 1/3$ for any direction, and for an infinite cylinder magnetized along the axis, $N = 0$. Analytic expressions for demagnetizing factors in uniformly magnetized ellipsoids and rectangular prisms are given in Dunlop and Özdemir (1997).

The dependence of the demagnetizing factor on the shape of particle produces a uniaxial anisotropy called shape anisotropy. The uniaxial anisotropy energy is given by

$$\begin{aligned} E_u &= K_u V \sin^2\theta \\ K_u &= \frac{1}{2}\mu_0 \Delta N M_S^2 \end{aligned} \qquad [15]$$

where θ is the angle between M and the easy direction ($\theta = 0°$, 180°), K_u is the uniaxial anisotropy constant, and ΔN is the difference in demagnetizing factors between the $\theta = 0°$ and $\theta = 90°$ directions. In strongly magnetic (high M_S) materials like magnetite and iron, shape anisotropy dominates over the intrinsic magnetocrystalline anisotropy for particle shapes that depart even slightly from ideal spherical or cubic shapes.

11.05.4.3.3 Single-domain, multidomain, and pseudo-single-domain particles

The main driving force for domain formation is the reduction of demagnetizing or magnetostatic self-energy (free pole effect) through various types of flux closure configurations, which, in sufficiently large particles, reduce energy enough to outweigh the added energy associated with forming domain walls. Domains reduce the net M and thereby lead to lower magnetostatic energy states relative to the uniform magnetization state. An SD grain is uniformly magnetized with M_S along an easy axis, thus minimizing exchange and anisotropy energies (Figure 7(a)). However, depending on grain volume, the magnetostatic energy of this configuration may not be minimal. Further reduction in magnetostatic energy can be gained if the grain divides into two domains each magnetized along the easy axis but with M_S in opposite directions (Figure 7(b)), thus reducing the remanent moment of the grain. This subdivision into more and more domains (Figure 7(c)) cannot continue indefinitely because the domain walls require additional exchange and magnetocrystalline energy. This comes about as the magnetization changes smoothly from that of one domain to the other, and spins within the wall are no longer parallel to each other or aligned along an easy direction. Eventually, an equilibrium number of domains will be reached, which reflects

the balance between magnetostatic, anisotropy, and domain wall energies for a given grain volume. Therefore, micromagnetic states are strongly volume-dependent, and as grain size increases, the number of domains or the complexity of flux closure features also increases. Experimental observations from direct imaging of surface domain structures confirm these theoretical predictions (e.g., Halgedahl, 2007).

The importance of domains in controlling bulk magnetic behavior is made apparent by the fundamentally different behaviors exhibited by SD and MD particles in response to an applied field. In SD particles, changes in magnetization can occur only by moment rotation over an anisotropy barrier from one easy axis to another. In contrast, changes in magnetization in MD particles occur initially by wall motion in low fields, as domains whose magnetizations are oriented closest to H grow at the expense of other domains, followed by moment rotation in higher fields, once the applied field is strong enough to eliminate all domain walls. Moment rotation is always a harder process, requiring larger fields to accomplish, than wall motion. In other words, coercivities for SD grains are always larger than for MD grains. Furthermore, unlike SD grains, MD grains are not uniformly magnetized except at saturation in a large field. For example, hysteresis loops for MD (3-mm-sized single crystal) and SD (50-nm-sized particles) magnetite are shown in **Figure 9**. The MD loop is narrow and ramp-like with low values of saturation remanence and coercivity, while the SD loop is much wider and rectangular-like with higher values of saturation remanence and coercivity. However, both loops are fully saturated in similar fields. For the MD sample, the initial slope ($1/N$, eqn [7] and **Section 11.05.4.3.2**) is nearly constant until the approach to saturation begins around 200 mT ($\sim$160 kA m^{-1} $\sim M_S/3 \sim NM_S$).

Theory predicts a dramatic drop in remanence and coercivity as grain size increases from SD to MD. However, only a gradual change in these properties is observed that extends well above the SD–MD transition size. For magnetite, remanence and coercivity vary approximately as d^{-n} ($n \sim 0.5$–0.7) for the size range 0.1–400 μm (Dunlop, 1986). The intermediate state between SD and MD behaviors is referred to as PSD. These grains are large enough to contain domains but small enough that the domain walls are not as mobile; therefore, PSD magnetization is more stable with respect to time, temperature, and field and it behaves magnetically more like SD grains than MD grains. The importance of PSD behavior in (titano)magnetite is that the grain size range for PSD behavior covers the range in sizes that most commonly occur in natural materials.

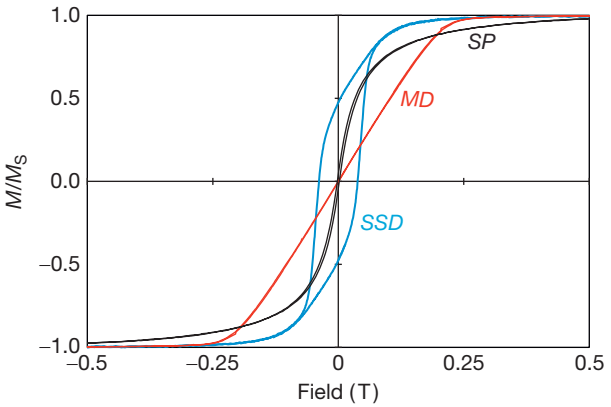


Figure 9 Examples of SSD (stable single domain), SPM (superparamagnetic), and MD (multidomain) hysteresis loops for magnetite. The SSD sample contains chains of $\sim$50 nm sized biogenic magnetite from magnetotactic bacteria, dispersed in a nonmagnetic matrix; the SPM sample contains $\sim$10 nm sized biogenic magnetite from iron-reducing bacteria, and the MD sample is a $\sim$3 mm diameter sphere from a synthetic single crystal of magnetite.

11.05.4.3.4 Critical sizes

As grain size decreases, a critical size (d_0) will be reached where the grain can no longer accommodate a domain wall because the energy cost to include a wall is much greater than the savings in magnetostatic energy achieved by subdividing into domains. Below this critical size, the SD structure is the lowest energy state. The critical size for the SD state, therefore, represents a balance between magnetostatic energy (depending on M_S and the shape of the grain) and wall energy (depending on the magnetocrystalline anisotropy (K_1), exchange (A), and magnetostriction constants). As a result, different minerals will have different critical sizes. A classic calculation for the critical size at which the energy of the SD and two-domain (TD) structures in a sphere (or cube) is just equal has been given by many authors (e.g., Kittel, 1949). This calculation yields the result

$$d_0 = \frac{a\gamma_w}{\mu_0 N_{SD} M_S^2} \qquad [16]$$

where $a = 6$ (sphere) or 4 (cube), γ_w = wall energy ($\sim\sqrt{KA}$), and N_{SD} = demagnetizing factor (Dunlop and Özdemir, 1997). The calculation assumes a thin planar wall such that the magnetostatic energy produced by surface termination of the wall is negligible. The usefulness of eqn [16], even with simplifying assumptions, is that it shows how d_0 is strongly dependent on M_S. Minerals with high M_S, like titanomagnetite or pyrrhotite, will have much smaller transition sizes than minerals with low M_S, like hematite. In addition, elongated grains ($N \sim 0$) will have larger transition sizes than more equant grains.

More rigorous 3D micromagnetic calculations allow other spin configurations with more efficient magnetostatic energy minimization to form (e.g., Dunlop and Özdemir, 1997; Harrison et al., 2007). In fact, the simple TD state may not be the lowest-energy ground-state configuration for grains sizes just above d_0. Micromagnetic calculations predict a variety of other microstates (flower and vortex states, **Figure 10**) with lower energy configurations than the simple uniform TD states (e.g., Fabian et al., 1996; Muxworthy et al., 2003; Williams and Dunlop, 1989). The single-vortex configuration (**Figure 10(b)**) resembles the magnetic structure in **Figure 7(d)** but has spin orientations that vary continuously through the grain, rather than changing abruptly across the walls that separate uniformly magnetized domains. Furthermore, these alternative microstates can be local energy minima (LEM) rather than global energy minima or ground states. This means that grains can be trapped in higher-energy LEM states until perturbed by magnetic fields, initiating the transformation to a lower energy state. An important LEM state is the metastable SD state, in

which the nearly uniformly magnetized SD state (flower state) metastably persists in larger-sized grains whose ground-state configurations are predicted to be a non-SD vortex configuration.

Metastable SD states and alternative LEM states for the same grain have been experimentally observed in magnetite, titanomagnetite, and pyrrhotite for grain sizes >1000 nm, well above their equilibrium values for d_0 (e.g., Dunlop and Özdemir, 2007; Halgedahl, 2007). For magnetite, the best estimate for the SD transition size in equant particles is about 80 nm. A summary of theoretical critical sizes for common magnetic minerals is given in Table 3.

11.05.4.3.5 Superparamagnetism

Another important critical transition size occurs when SD grains are small enough to exhibit superparamagnetism and particles lose their coercivity and their ability to retain remanence. An SD particle of volume V has a uniform magnetization aligned along an axis where the anisotropy energy is a local minimum. If V is small enough or temperature high enough, thermal fluctuations will be sufficient to cause spontaneous transitions of the magnetization over the anisotropy barrier even when $H=0$. When this happens, the grain becomes SPM, M_R and H_C go to zero, and the M–H curves become completely reversible (Figure 9). This transition therefore separates unstable SPM behavior from stable SD (SSD) behavior. For magnetite, the SPM transition at room temperature occurs at a grain size of about 30 nm (Dunlop and Özdemir, 1997). Magnetic relaxation and superparamagnetism are responsible for magnetic viscosity and frequency-dependent susceptibility. These concepts are developed further in the next section.

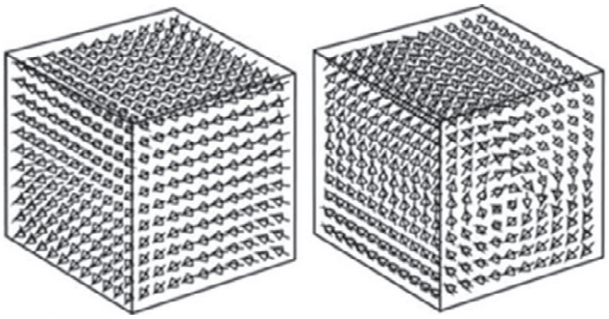


Figure 10 Numerical micromagnetic results for a cubic grain of magnetite (120 nm) at room temperature showing two different possible local energy minima (LEM) states: (a) flower state, (b) vortex state. Figure modified from Muxworthy AR, Dunlop DJ, and Williams W (2003) High-temperature magnetic stability of small magnetite particles. *Journal of Geophysical Research* 108 (B5): EPM 18-1–EPM 18-9, http://dx.doi.org/10.1029/2002JB002195.

Table 3 Critical transition sizes in common magnetic minerals at room temperature

Mineral	*Superparamagnetic size (nm)*	*Single-domain size (nm)*
Magnetite	25–30	50–84
Maghemite	20–30	60
Titanomagnetite ($x\sim0.6$, $z=0$)	80	200–600
Titanomaghemite ($x=0.6$, $z=0.7$)	90	2400
Hematite	25–30	15 000
Pyrrhotite	n.d.	1600
Greigite	17	200
Iron	8–26	17–23

Data are a compilation of experimental and numerical results mainly for equidimensional-sized particles from Dunlop and Özdemir (1997, 2007), Roberts et al. (2011), and Moskowitz et al. (1997).

11.05.4.4 Thermal Relaxation and Néel Theory

The magnetostatic energy of a magnetic dipole in an external magnetic field (the Zeeman energy, eqn [3]) is minimized when the dipole moment is parallel to the field. In fine-particle assemblages, the alignment of magnetic moments with an external field is impeded by energy barriers due to shape and magnetoelastic and/or magnetocrystalline anisotropy, and it is also disrupted by the randomizing effects of thermal energy. For all particle sizes and compositions, there is a tendency for the net moment of a population of magnetic particles to evolve toward alignment with an external field, but a rigorous quantitative physical theory exists only for single-domain particles with uniaxial anisotropy (i.e., with two antiparallel preferred orientations for the magnetic moment).

Néel theory (Néel, 1949, 1955) describes the evolution of the magnetization of assemblages of SPM–SSD particles, as functions of time, temperature, and ambient magnetic field. It is important to recognize that in the equilibrium state, the particle moments are generally not all in the minimum-energy orientation; a significant fraction may be aligned antiparallel to the external field, kicked into the higher-energy state by random thermal fluctuations. The proportions in high- and low-energy orientations follow the Maxwell–Boltzmann distribution, $N(E)\propto\exp\left(-\frac{E}{kT}\right)$.The equilibrium magnetization M_{eq} in an applied field H for an aligned assemblage of uniaxial SPM or SSD particles is thus

$$M_{eq}(V,M_S,H,T) = M_S\tanh(\alpha) = M_S\tanh\left(\frac{\mu_0 V M_S(T) H}{kT}\right) \quad [17]$$

At low temperature and in strong fields, moments are strongly aligned and M_{eq} approaches M_S. As T approaches T_C, M_S decreases rapidly, moment orientations become strongly randomized, and M_{eq} approaches zero. For a randomly oriented assemblage, the equilibrium magnetization is described by the Langevin function $L(\alpha)$:

$$M_{eq} = M_S L(\alpha) = M_S[\coth(\alpha) - 1/\alpha] \quad [18]$$

which for small α approximates to $\alpha/3 = m_0 V M_S^2/(3kT)$, one-third the value for the aligned assemblage.Magnetization evolves toward equilibrium according to the rate law

$$dM/dt = (M_{eq} - M)/\tau \quad [19]$$

where the relaxation time in weak fields is

$$\tau(V,M_S,H_k,T) = \tau_0\exp\left(\frac{V M_S(T) H_k(T)}{2kT}\right) \quad [20]$$

The constant τ_0 is an atomic reorganization time related to lattice vibrations and is $\approx 10^{-9}$ s for FM and FI materials (e.g., Moskowitz et al., 1997) and H_k is the microscopic coercivity.

11.05.4.4.1 DC magnetization

If the external field and temperature do not vary with time, M_{eq} is constant, and the solution to the differential eqn [19] is

$$M(t) = M_0 e^{-t/\tau} + M_{eq}\left(1 - e^{-t/\tau}\right) \quad [21]$$

where M_0 is the magnetization at $t=0$. When τ is short compared to the measurement timescale τ_m or some other timescales of interest, equilibration occurs 'quickly,' and the particles are considered SPM or 'unblocked.' The SSD state is defined by $\tau \gg \tau_m$, and the transitional region where $\tau \sim \tau_m$ is characterized by readily observable time dependence; this is referred to as magnetic viscosity.

In zero field, the equilibrium magnetization is also zero, and at constant temperature, a remanent magnetization decays exponentially with time, the rate being governed primarily by V and H_k. Above a threshold size or blocking volume $V_b(T, H_k, \tau_m)$, isothermal decay is negligible and the magnetization is 'blocked.' Stability with respect to temperature is a key concept for remanent magnetizations. For an ensemble of identical particles with volume v, the relaxation time decreases sharply with increasing temperature, and at the blocking temperature $T_b(V, H_k, \tau_m)$, $\tau \sim \tau_m$. Above T_b, equilibration is effectively instantaneous.

11.05.4.4.2 AC susceptibility

In thermally stable materials, low-field susceptibility is due to small reversible rotations of magnetic moments away from easy axes toward the applied field and/or slight reversible displacements of domain walls from their zero-field energy-minimum locations. For nanoparticles with sizes at and below the SPM–SSD boundary, thermal fluctuations allow moments to rotate over much larger angles to align with an external field, 'jumping' over intervening maxima in anisotropy energy. This thermally activated irreversible susceptibility component may be much larger than the reversible susceptibility of thermally stable grains.

In an applied AC field of the form $H(t) = H_0\cos(\omega t)$ (where $\omega = 2\pi$ times the frequency f), the equilibrium magnetization in eqn [18] also varies with time, and the solution to eqn [19] is

$$M(t) = \frac{M_{eq0}}{(1+\omega^2\tau^2)}[\sin(\omega t) - \omega\tau\cos(\omega t)] \quad [22]$$

where M_{eq0} is the equilibrium magnetization in H_0. From this, we obtain the in-phase susceptibility

$$k' = \frac{M_{eq0}}{H_0(1+\omega^2\tau^2)} \quad [23]$$

and the out-of-phase (aka quadrature or imaginary) component

$$k'' = \frac{\omega\tau M_{eq0}}{H_0(1+\omega^2\tau^2)} \quad [24]$$

Equations [23] and [24] describe the thermally activated component of susceptibility, that is, the part due to irreversible rotation of moments over anisotropy energy barriers into minimum-energy orientations along the applied field. These components vanish for sufficiently high frequencies and/or long relaxation times ($\omega\tau \gg 1$), that is, for the SSD state. The quadrature component also vanishes for sufficiently low frequencies and/or short relaxation times ($\omega\tau \ll 1$), that is, the SPM state, where k' becomes equal to the equilibrium (DC) susceptibility $k_{sp} = M_{eq}/H_0 \simeq \mu_0 V M_S^2/(3k_B T)$.

11.05.4.5 Remanence Acquisition Processes

11.05.4.5.1 Natural remanent magnetization

NRM is a remanent magnetization acquired by rocks, sediments, soils, and archaeological materials in the geomagnetic field through various geologic processes and before any laboratory demagnetization or remagnetization treatment. Also, iron objects related to human activities buried in the near-subsurface can acquire a remanent magnetization during or after manufacturing (e.g., steel drums, pipelines, and UXO). The NRM is a vector composed of one or more remanent components acquired by different subpopulations of magnetic grains at various times during the rock's geologic history. The most important is the primary magnetization acquired at or near the time of rock formation that can persist over geologic time. All other components acquired after formation and some at much later times are called secondary magnetizations. The NRM is usually acquired parallel to the Earth's magnetic field at the time of formation or alteration. In some cases, the secondary magnetization process may totally remagnetize and replace the primary NRM, while in other cases, it may just mask or overprint it, leaving a vestige of the primary signal in some fraction of the remanence-carrying population of mineral grains. Secondary magnetizations can be acquired during subsequent chemical changes of the iron mineralogy during weathering, diagenesis, or metamorphism, or from exposure to magnetic fields from lightning strikes and drilling, or from the long (millions of years) exposure to the geomagnetic field at ambient or elevated temperatures. The great successes of paleomagnetism have come from laboratory techniques of partial demagnetization and statistical tools that can progressively remove secondary components and isolate and retrieve the primary NRM (e.g., Dunlop and Özdemir, 1997; Tauxe, 2010; Turner et al., 2007). However, in terms of magnetic anomaly characteristics, the total NRM (magnitude and direction) is a key property for correct interpretation of magnetic anomalies. In some cases, the remanent magnetization can be much larger than the induced magnetization and have directions entirely different from the present-day field.

NRM intensity is one of the most variable of magnetic parameters because it depends not only on the applied field but also on rock type, mineralogy, grain size, mode of remanence acquisition, and thermal and magnetic history. There are many different types of primary and secondary remanences (**Table 4**) depending on how and when ferromagnetic grains acquire magnetization in the presence of the geomagnetic field or other external fields.

Magnetization responsible for magnetic anomalies is the vector sum of the NRM and the induced magnetization. For most types of rocks, the induced magnetization is in the direction of the present-day geomagnetic field. The relative

Table 4 Major types of remanent magnetization (RM)

Remanent magnetization		*Blocking process*
Anhysteretic remanent magnetization	ARM	Combined effects of a large alternating field and a weak DC field at constant T
Chemical remanent magnetization	CRM	Growth/alteration of magnetic grains in weak H_a below T_c
Detrital remanent magnetization	DRM	Deposition of magnetic grains in weak H_a
Isothermal magnetization	IRM	Instantaneous exposure to large H_a at constant T
Natural remanent magnetization	NRM	Acquired by a rock through natural processes as opposed to artificial remanence produced in laboratory
Piezoremanent magnetization	PRM	Stress-induced magnetization in weak H_a at constant T
Partial TRM	pTRM	Partial TRM acquire by reheating to a temperature below T_c and cooling in weak H_a
Saturation remanent magnetization	SIRM	Instantaneous exposure to large H_a at constant T to saturate magnetization
Thermoremanent magnetization	TRM	Cooling through T_c in weak H_a
Thermoviscous magnetization	TVRM	Long-term exposure to weak H_a at elevated T
Viscous remanent magnetization	VRM	Long-term exposure to weak H_a at constant T

T_c, Curie temperature; H_a, applied field.

importance of NRM compared with induced magnetization is characterized by the Koenigsberger ratio Q_n, a dimensionless quantity given by

$$Q_n = \frac{J_{nrm}}{k_i H_e} \quad [25]$$

where J_{nrm} is the magnitude of the NRM (per unit volume), k_i is the volume susceptibility, and H_e is the magnitude of the Earth's magnetic field at the site under consideration. Remanence-dominated anomalies are characterized by $Q_n > 1$ and in some cases can lead to negative anomalies when the NRM direction is opposite to the present-day field direction. For example, a large negative anomaly of ~13 000 nT is associated with a norite unit in the Bjerkreim–Sokndal layered intrusion in Norway (**Figure 5**) containing exsolved hemoilmenite with high NRM values (1–74 $A\,m^{-1}$), negative inclinations, and Q_n values (>5) (McEnroe et al., 2004).

11.05.4.5.1.1 Thermoremanent magnetization

TRM is the remanence acquired by magnetic grains as they cool from above the Curie temperature to room temperature in the presence of a weak magnetic field (usually 50–100 μT). It is the typical mode of remanence acquisition in igneous rocks (extrusive and intrusive), high-grade metamorphic rocks, and some archaeological materials (e.g., fired ceramics, bricks, hearths, and baked soils). TRM is usually a primary remanence unless the sample is reheated above its maximum Curie temperature. The theory for TRM in SD grains was proposed by Néel (1949, 1955), and in essence, it involves the evolution of the magnetization of a population of SD particles on cooling from T_C, remaining in equilibrium with the external field (eqn [18]) as long as the relaxation time (eqn [20]) is sufficiently short compared to the cooling time and eventually becoming blocked in or stabilized when τ becomes long. Once the population is blocked in (at $T < T_C$), the orientation distribution of particle moments is fixed, and on continued cooling, the only change is increasing intensity due to the temperature-dependent spontaneous magnetization of the magnetic material. A detailed review of Néel theory is given by Dunlop and Özdemir (1997).

Anomaly sources produced by induced magnetization in the geomagnetic field reflect variations in the abundance of high-susceptibility magnetic phases, which are usually coarse-grained (>1–10 μm) and in the MD state. Remanence-dominated anomalies, while generally reflecting modal abundances of magnetic phases, are also controlled by the mode of remanence acquisition (e.g., TRM, chemical remanent magnetization (CRM), or viscous magnetization (VRM)) and magnetic domain state (SD vs. MD). The most intense natural remanence (per unit volume) acquired in weak magnetic fields like the geomagnetic field is TRM carried by the fine-grained (<1 μm) fraction in the SD size range (Dunlop and Özdemir, 1997). As grain size increases above the SD transition size (<0.1 μm for magnetite and <2.0 μm for titanomagnetite and pyrrhotite), MD micromagnetic states develop with a concomitant decrease in TRM intensity. For example, the intensity of weak-field TRM in magnetite decreases by nearly two orders of magnitude as grain size increases from 0.1 to 10 μm (**Figure 11**).

Surprisingly, in hematite, TRM intensity is actually larger in coarse-grained MD particles (>1–1000 μm) than in fine-grained SD particles (**Figure 11**), a TRM–grain size relationship that is exactly opposite to that found in magnetite and titanomagnetite (Dekkers and Linssen, 1989; Dunlop and Kletetschka, 2001; Kletetschka et al., 2000a,b). What may be even more surprising is that, as Kletetschka et al. (2000a,b) found, the intensity of weak-field TRM (per unit volume) in MD hematite can be greater than thermoremanences in other typical MD minerals (magnetite, titanomagnetite, and pyrrhotite), and millimeter-sized hematite can even overlap the low-end ranges of TRM in submicron magnetite and nominally SD pyrrhotite. This is even more remarkable when one considers that the saturation magnetization (M_S) in magnetite is 200 times larger than for hematite. These observations have elevated the status of coarse-grained MD hematite occurring in igneous and metamorphic rocks as a potential source for remanence-dominated anomalies on Earth and Mars (Dunlop and Arkani-Hamed, 2005; Dunlop and Kletetschka, 2001). Ti-poor titanohematite may also display the same TRM–grain size trend as hematite described earlier, but there are no data to confirm this trend.

Hematite has a canted antiferromagnetic spin structure with a weak spontaneous magnetization. Among other things, the low M_S compared to magnetite makes the demagnetizing

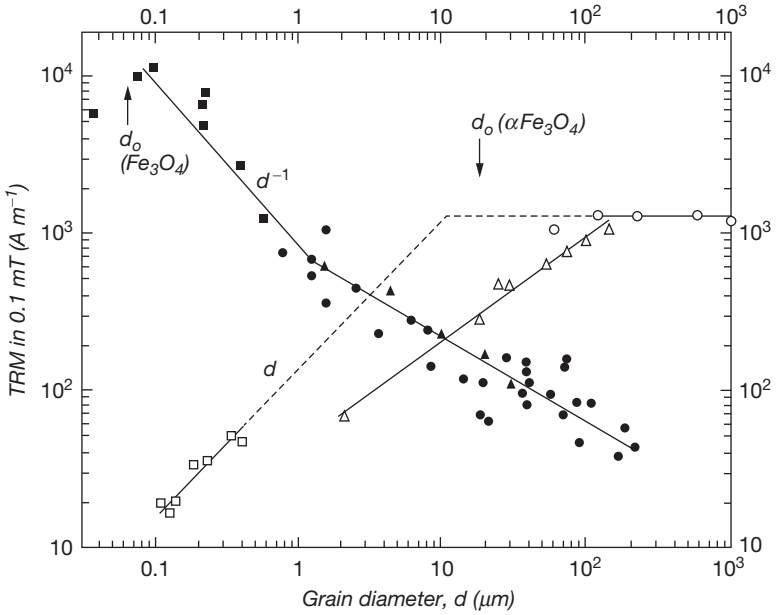


Figure 11 Dependence of thermoremanent magnetization (TRM) intensity on grain size for magnetite (solid symbols) and hematite (open symbols) from various sources (see Dunlop and Arkani-Hamed (2005) for more details). The single-domain transition size (d_0) is shown for each mineral above which TRM varies as $\sim d^{-1}$ for magnetite and $\sim d$ for hematite. Figure from Dunlop DJ and Arkani-Hamed J (2005) Magnetic minerals in the Martian crust. *Journal of Geophysical Research* 110: E12S04. http://dx.doi.org/10.1029/2005JE002404.

energy (eqn [14]), which drives domain walls backs toward their equilibrium, demagnetized (zero remanence) states, much less important in hematite than in magnetite. Whereas in MD magnetite, H_d limits the intensity of remanence to a very small fraction of M_S, this is apparently not the case for hematite where walls can be trapped in LEM far from the demagnetized state in a 'high' remanence state.

11.05.4.5.1.2 Depositional remanent magnetization

The geomagnetic field, though weak, exerts a significant torque (eqn [4]) on a detrital particle having a stable magnetic moment. Such particles therefore tend to become oriented with their moments aligned along the ambient field direction, and despite the complexity of processes involved in sedimentation and postdepositional consolidation, the net primary magnetization of sediments and sedimentary rocks generally parallels the depositional field orientation. The efficiency of alignment is low, since gravitational, hydrodynamic, and bioturbational mechanisms generally exert stronger controls on particle orientations than the magnetic field (e.g., Tauxe and Yamazaki, 2007). A consequence of the low alignment efficiency is that the remanent magnetization of sediments and sedimentary rocks can generally be neglected in crustal magnetic field modeling, being typically much less than the magnetization induced *in situ* by the geomagnetic field (see **Section 11.05.5.2**). However, exceptions exist and include certain clastic sedimentary rocks containing detrital magnetic minerals (Grauch and Hudson, 2011; Reynolds et al., 1990) and the intense anomalies associated with banded iron formations.

Depositional remanent magnetization (DRM) theory involves pre-, syn- and postdepositional mechanisms that act to orient particles according to their shapes and/or their magnetic moments. Magnetic particles may remain mobile and able to reorient within sediment pore spaces after deposition, until dewatering, compaction, and consolidation/cementation finally 'lock in' the particle orientations and stabilize the DRM (known as postdepositional RM or pDRM). Locking in of remanence at some depth below the sediment–water interface (the 'lock-in depth') introduces a time lag into the recording of ambient field variations (e.g., Denham and Chave, 1982; Liu et al., 2008; Tarduno and Wilkison, 1996), and locking in over a range of depths results in a smoothing or blurring of the paleofield record. There are many complexities in both DRM and pDRM processes and a detailed treatment is beyond the scope of this article, but recent overviews are available (Bilardello et al., 2013; Mitra and Tauxe, 2009; Shcherbakov and Sycheva, 2010; Tauxe et al., 2006). Areas of continuing research include inclination shallowing, that is, a bias of the DRM orientation toward the horizontal due to gravitational torques and postdepositional compaction (Bilardello and Kodama, 2010; Bilardello et al., 2013; Tan et al., 2007; Tauxe, 2005; Tauxe et al., 2008); factors governing the relationship between DRM intensity and ambient field strength during deposition and lock-in (Hofmann and Fabian, 2009; Yamazaki, 2008); and flocculation, that is, mutual attraction and coagulation of magnetic and nonmagnetic particles such as clays, due to electrostatic and van der Waals forces, and its effects on DRM orientation and intensity (Mitra and Tauxe, 2009; Shcherbakov and Sycheva, 2010; Tauxe et al., 2006).

11.05.4.5.1.3 Chemical remanent magnetization

In TRM and DRM, the crystal chemistry and particle sizes of the remanence carriers are constant, and only the orientations of the particles or of their magnetic moments change. Many surficial processes result in changes of bulk chemical

composition or chemical zonation of primary iron-bearing minerals or formation of new minerals, and these transformations commonly result in the acquisition of a chemical remanent magnetization (CRM).

When magnetic minerals form by nucleation and growth, they are initially in an SPM state, in equilibrium with the ambient field, and with continued growth, they pass through the blocking volume (**Section 11.05.4.4.1**), freezing the magnetic orientations in the SSD state. Remanence formed in this way is referred to as 'growth CRM,' which has some similarity to the TRM process (e.g., Fabian, 2009; McClelland, 1996). If particle volumes do not grow beyond the upper critical size for SD behavior, the resultant CRM intensity (per unit mass of remanence-carrying material) is comparable to that of TRM.

Alteration of preexisting magnetic minerals may occur by a wide variety of mechanisms including oxidation/reduction (involving changes in bulk chemical composition, e.g., oxidation of magnetite to maghemite or hematite), inversion (involving changes in crystal structure, e.g., inversion of maghemite to hematite), subsolvus unmixing of metastable homogeneous phases (e.g., formation of hematite/ilmenite or magnetite/ulvöspinel intergrowths), and oxyexsolution (e.g., formation of ilmenite lamellae in titanomagnetites). These result in what are often called 'parent–daughter' CRMs, and much research has been done on the extent to which these are controlled by the contemporaneous ambient field and by the primary remanence of the parent phase (Gendler et al., 2005; Heider and Dunlop, 1987; Kelso et al., 1991; Özdemir and Dunlop, 1985, 1993). In general, daughter CRMs are less intense than parent TRM (e.g., Fuller et al., 2002) although under some conditions, the reverse may occur (e.g., Fabian, 2009). An important example where the CRM direction is inherited from the parent TRM is found in the ocean crust where low-temperature maghemitization of primary titanomagnetite ($x \sim 0.6$) in MORB still carries the original polarity reversal record formed at mid-ocean ridges.

11.05.4.5.1.4 Viscous and thermoviscous remanent magnetization

VRM is the magnetization acquired at constant temperature and pressure and without any chemical changes, during long exposure to weak magnetic fields (< 100 μT), like the geomagnetic field. A complementary process is the viscous decay of magnetization in zero magnetic field. VRM is a consequence of the thermally activated nature of magnetization given by Néel theory (eqns [17]–[21]), which describes the combined effects of geologic time and elevated temperatures (Dunlop and Özdemir, 1997; Néel, 1949). VRM is a secondary remanence. Once an initial magnetization is blocked in (e.g., TRM, DRM, or CRM) at the time of formation, magnetic grains with short relaxation times ($\tau <$ age of the rock) will unblock and slowly reequilibrate to the applied field over the long geologic history of the rock. This usually means that VRM is parallel to present-day field acquired over the Brunhes polarity chron (the past 0.78 Ma); therefore, VRM can be assumed to be codirectional with the induced magnetization.

The intensity of magnetization increases (acquisition in an applied field) or decreases (decay in zero field) with the logarithm of time, J_{vrm}, $\propto S \log t$, although experimentally, it is found that the acquisition process is often nonlinear in $\log(t)$ (Dunlop, 1983; Yu and Tauxe, 2006). The slope of this curve ($S = \mathrm{d}(J_{vrm})/\mathrm{d}(\log t)$) is characteristic of the relaxation spectrum and is usually referred to as the viscosity coefficient. This coefficient is a function of H and T, and VRM acquisition increases with both applied field strength and at elevated temperatures (Dunlop, 1983; Williams and Muxworthy, 2006). VRM typically refers to magnetization acquired at ambient temperature, while thermoviscous remanent magnetization (TVRM) is a VRM acquired at elevated temperatures (Pullaiah et al., 1975). In the latter case, elevated temperature over extended time periods enhances the acquisition process and produces a TVRM that may eventually replace the original NRM.

Extremely fine grains near the SP–SD transition threshold (< 20 nm) and coarser PSD–MD grains (> 1–10 μm) are susceptible to viscous remagnetization of their primary NRM. While the theory of VRM in MD particles is incomplete, the mechanism is likely related to thermal activation of wall segments over energy barriers. VRM carried by MD grains can be significant. For example, Kelso and Banerjee (1994) found that deep-crustal mafic granulites of Central Australia containing MD magnetite can rapidly acquire VRM when heated above 200 °C, and when extrapolated over the Brunhes epoch, the VRM can entirely overprint the original NRM and produce magnetizations of ~ 4 A m^{-1}.

11.05.4.5.2 Artificial remanent magnetization

Besides laboratory-based susceptibility measurements, several artificial remanent magnetizations acquired under laboratory conditions of controlled field, temperature, or time are used to characterize magnetic materials. Along with susceptibility and hysteresis, artificial remanent magnetization forms the tool kit of magnetic methods used in environmental magnetism, which has made numerous contributions to studies of climate, sedimentology, oceanography, archaeology, and anthropogenic pollution over the past 25 years (e.g., Evans and Heller, 2003; Liu et al., 2012; Maher, 2011; Thompson and Oldfield, 1986). Remanent magnetizations imparted in the laboratory are useful for material characterization, especially through their intensities (related to the compositions, concentrations, and grain size distributions of magnetic minerals); their stability spectra, that is, coercivity and unblocking-temperature spectra (related to compositions and particle size distributions of the magnetic minerals); and their anisotropy (related to depositional, emplacement, and/or deformational fabrics). The two most common types of artificial remanent magnetization are anhysteretic remanent magnetization (ARM) and isothermal remanent magnetization (IRM).

11.05.4.5.2.1 Anhysteretic remanent magnetization

ARM is the magnetization acquired at constant temperature when a sample is subjected to an AC field of gradually decreasing amplitude (H_{AF}) simultaneously with a steady, unidirectional DC field (h_{dc}). For isotropic samples, the direction of the ARM is parallel to h_{dc} but the intensity will depend on the amplitudes and relative orientations of the AC and DC fields. The maximum ARM intensity occurs when H_{AF} and h_{dc} are parallel. Typical experimental protocols use AC fields of 100–200 mT and DC fields of 0.05–0.1 mT (i.e., h_{dc} is a weak field like the geomagnetic field). ARM is not a naturally occurring remanence but it is used in a variety of applications in

paleomagnetism. For example, ARM is used to (1) estimate absolute paleointensity from igneous rocks and relative paleointensity from lake and marine sediments, (2) characterize magnetic carriers and determine domain state and grain size, and (3) detect magnetic fabrics in rocks and sediments. ARM in combination with susceptibility of soils has been used to map culturally modified landscapes associated with archaeological sites (**Section 11.05.5.4.3**). More detailed reviews of ARM can be found in Egli (2006) and Moskowitz (2007).

11.05.4.5.2.2 Isothermal remanent magnetization

Although other remanence-acquisition processes commonly occur isothermally (e.g., ARM and CRM), the term IRM is generally understood to involve the application and removal of a significant DC magnetic field, that is, one approaching the coercivity of ferrimagnetic materials (on the order of several millitesla to tens or hundreds of millitesla). Such intense fields are rare in natural terrestrial environments, but IRMs can be imparted by lightning strikes (e.g., Maki, 2005; Shimizu et al., 2007; Verrier and Rochette, 2002). In a sufficiently large field to cause saturation, the resulting IRM is called a saturation remanent magnetization (SIRM). Most commonly, IRMs are imprinted in the laboratory for the same purposes listed earlier for ARMs. In addition, detailed stepwise demagnetization of strong-field IRM (as well as ARM) provides data sets that can be used to analyze the coercivity spectrum and identify and characterize component populations of remanence-carrying particles having geologically meaningful commonalities of mineralogy and particle size (Egli, 2004a,b,c; Heslop and Dillon, 2007; Heslop et al., 2002). One last note on nomenclature is necessary. IRM is related to hysteresis and common practice uses M_R for the IRM determined by measurement of a hysteresis loop (and M_{RS} if saturation is achieved), whereas IRM and SIRM are typically used for remanence measurements made with zero-field magnetometers after magnetization on a separate instrument (electromagnet, pulse magnetizer, etc.).

11.05.4.6 Complex Susceptibility and Frequency Dependence of Susceptibility

Equation [22] shows how the magnetization in a sinusoidally time-varying field can, through the mechanism of magnetic viscosity, have both in-phase and out-of-phase components, that is, it can lag behind the driving field. The out-of-phase component is also known as the imaginary or quadrature susceptibility, and in addition to viscosity, it may be produced by two other mechanisms: electrical conductivity and low-field hysteresis, both of which will be described later. In general, an externally applied time-varying magnetic field $H(t)=H_0\cos(\omega t)$, where H_0 is the amplitude and ω the angular frequency, results in a quasi-sinusoidal magnetization that may lag behind the time-varying applied field $M(t)=M_0\cos[\omega(t-\Delta t)]=M_0\cos(\omega t-\delta)$, where Δt is the time lag and δ is referred to as the phase shift or lag (or simply as the phase). Mathematically, this is identical to the sum of an in-phase variation $M'\cos(\omega t)$ and an out-of-phase variation $M''\sin(\omega t)$, with $M'^2+M''^2=M_0^2$ and $\tan(\delta)=M''/M'$. In addition, $M(t)$ may contain higher harmonics.

Conductive effects are generally negligible for most geologic materials, the exceptions being mainly sulfide minerals and graphite. A magnetic field changing with time generates an electric field and thus, in conductive material, produces electric currents. The electric field intensity and current strength vary in proportion to the rate of change in the magnetic field (Faraday's law of induction), and thus, they are out-of phase with an applied AC field. The induced currents in turn produce secondary magnetic fields (Ampère's law), which are also out-of phase with the applied field, that is, if $H_{applied}$ varies as $\cos(\omega t)$, then $H_{secondary}$ varies according to $\sin(\omega t)$ or equivalently, as $\cos(\omega t-\pi/2)$. The time-varying secondary magnetic fields also generate currents and thus produce tertiary magnetic fields varying as $\cos(\omega t-\pi)$, or equivalently as $-\cos(\omega t)$, that is, a negative in-phase response. Moreover, the eddy currents dissipate energy, and the external field variations are consequently attenuated in the material, with a scale length ('skin depth' d) related to electrical conductivity and magnetic permeability of the material as well as frequency of the applied field, and response is thus a strong function of the size and shape of a conductive body. For a conductive sphere of radius R, conductivity σ, and permeability μ, in an AC field of angular frequency ω,

$$k'+ik'' = \frac{3\mu[\sinh\alpha-\alpha\cosh\alpha]+3\mu_0[\sinh\alpha-\alpha\cosh\alpha+\alpha^2\sinh\alpha]/2}{\mu[\sinh\alpha-\alpha\cosh\alpha]-\mu_0[\sinh\alpha-\alpha\cosh\alpha+\alpha^2\sinh\alpha]} \quad [26]$$

(Worm et al., 1993), where $\alpha=R(i\mu\sigma\omega)^{1/2}$. In-phase susceptibility decreases monotonically with increasing frequency for conductive materials, whereas the out-of-phase response increases to a peak value and then decreases (**Figure 12**). The eddy current response depends on the size of an object as well as its conductivity, and in general, the laboratory instruments used in rock magnetic research do not reach high enough frequencies to allow observation of a strong eddy current response from conductive particles in a resistive rock matrix (**Figure 12**). However, in electromagnetic geophysical exploration, this is the principal signal source.

Weak-field hysteresis was first measured and described by Rayleigh (1887), who found that the function $M(H)$ was essentially linear in samples of iron in DC fields up to about 5 mT ($\sim$4000 A m^{-1}) but, in higher fields, nonlinear behavior became significant in the initial magnetization curve ($M=\alpha H+\beta H^2$; $k=\alpha+\beta H$). In AC fields above a threshold strength, magnetization is both nonlinear and irreversible, that is, it follows minor hysteresis loops. Rayleigh calculated the Fourier series for parabolic minor loops in an AC field of amplitude H_0:

$$M(t)=\left[\alpha H_0+\beta H_0^2\right]\cos(\omega t)+\frac{4\beta H_0^2}{3\pi}\sin(\omega t) \quad [27]$$

(plus higher odd harmonics related to the waveform asymmetry). Thus, the magnetization time series is resolved into in-phase and quadrature components, and the corresponding susceptibilities are $k'=M'/H_0=\alpha+\beta H_0$ and $k''=4\beta H_0/(3\pi)$ components (both of which depend linearly on amplitude).

The threshold of $M(H)$ nonlinearity varies strongly with mineralogy and grain size. Magnetite generally behaves linearly in fields up to 1000 A m^{-1} or more. The onset of low-field hysteresis occurs at much lower fields (10–100 A m^{-1}) for such

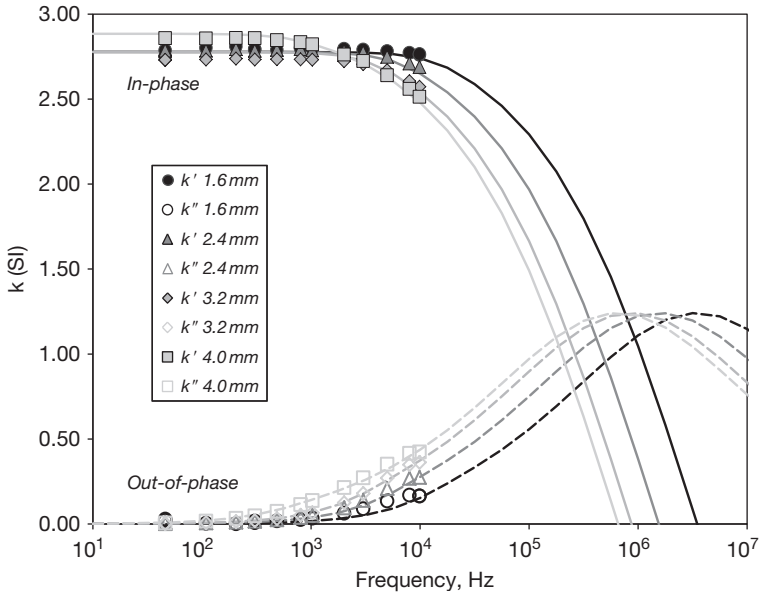


Figure 12 In-phase (k') and out-of-phase (k'') susceptibilities measured on steel spheres of different diameters at room temperature and at frequencies up to 10 kHz, using a Magnon Instruments VFSM susceptibility instrument. Solid (open) symbols indicate in-phase (out-of-phase) susceptibility. Data show significant magnetic and conductive (eddy current) contributions. Solid (in-phase) and dashed (out-of-phase) curves show values calculated from eqn [26], for frequencies up to 10 MHz.

important minerals as pyrrhotite, titanomagnetite, and hematite (Hrouda, 2002; Jackson et al., 1998; Worm, 1991; Worm et al., 1993), all of which generate significant quadrature signals in the AC fields (100–1000 A m^{-1}) used in commercial susceptibility instruments (**Figure 13**), but which generally do not closely adhere to Rayleigh's law (eqn [27]).

Viscous relaxation on millisecond to second timescales is an important control on AC susceptibility in the 1 Hz to several kHz frequency range. At room temperature, for sources consisting of magnetite or maghemite particles, such relaxation times roughly correspond to grain sizes in the 10–25 nm range. For low frequencies and/or short relaxation times ($\omega\tau << 1$), the assemblage is SPM and M stays in phase with and in equilibrium with the changing field: $M(t) = k_{SP}H(t) = M_{eq}(H_0)H(t)/(H_0)$. With increasing frequency and/or relaxation time, the time constants for measurement and relaxation become comparable ($\tau/\tau_m \sim \omega\tau \sim 1$), and the magnetization becomes more 'viscous,' lagging significantly, with roughly equal in-phase and quadrature components: $M' \sim M'' \sim M_{eq}(H_0)/2$. Finally, for very high ω and/or very long τ, the grains reach the SSD state ($k_{//} \to 0$). These relationships are illustrated in **Figure 14** for an ash-flow tuff sample from Yucca Mt. (Nevada, USA); these have very narrow size distributions, ranging from fine SPM at the flow margins to SSD in the flow interior; for the sample shown, the mean particle length is 20–25 nm (Schlinger et al., 1991; Till et al., 2010; Worm and Jackson, 1999). At 20 K, it is in the SSD state, and at room temperature, it is fully unblocked; in both cases, there is no significant out-of-phase signal nor is there any significant frequency dependence of the in-phase susceptibility. At intermediate temperatures, viscous superparamagnetism dominates the AC response: χ' rises smoothly from the SSD to the equilibrium SPM value, rapidly for low frequencies, and more slowly for higher f. Peaks of χ'' occur at $T_b(f,V)$, where χ' has approximately half its equilibrium value. However, χ'' is much less than χ' at $T_b(f,V)$, contrary to expectations (eqns [23] and [24] with $\omega\tau = 1$) for ensembles of identical particles; the discrepancy is due to the nonzero width of the particle size distribution in these samples. At higher temperatures above this unblocking interval ($\omega\tau \ll 1$), the particles are in a purely SPM state, $\chi' = \chi_{sp} \sim VM_S^2/(3kT)$, independent of frequency, and $\chi'' = 0$.

For a distribution of grain volumes $N(V)$ at constant T, there are corresponding distributions of equilibrium susceptibility $\chi_{eq}(V)$ and relaxation time $\tau(V)$, and eqns [23] and [24] become integrals over some range of V. The net effect is to broaden the peak in $\chi''(f)$ and diminish its amplitude relative to χ' or to χ_{eq}. Néel (1949) showed that when $N(V)$ is reasonably smooth near V_b, eqns [23] and [24] can be integrated and manipulated to yield

$$\chi'' = -\pi/2 \times d\chi'/d(\ln(f)) \qquad [28]$$

(see also Shcherbakov and Fabian, 2005).

Pedogenic particles of maghemite or magnetite close to the SPM–SSD transition size (~20–40 nm) are commonly found in many soil types and exhibit frequency-dependent susceptibility and magnetic viscosity (Dearing et al., 1996b; Evans and Heller, 2003). This type of behavior can adversely affect the performance of electromagnetic induction (EMI) detectors (frequency- or time-domain systems) commonly employed to search for buried UXO and land mines (Dabas and Skinner, 1993; Das, 2006; Hannam and Dearing, 2008; Igel et al., 2012). In general, the absolute values of susceptibility affect frequency-domain metal detectors, whereas the frequency dependence of susceptibility affects time-domain detectors. Magnetic viscosity of soils can degrade or reduce the sensitivity of EMI detectors because time-dependent magnetization changes can mimic eddy current decay in metallic objects

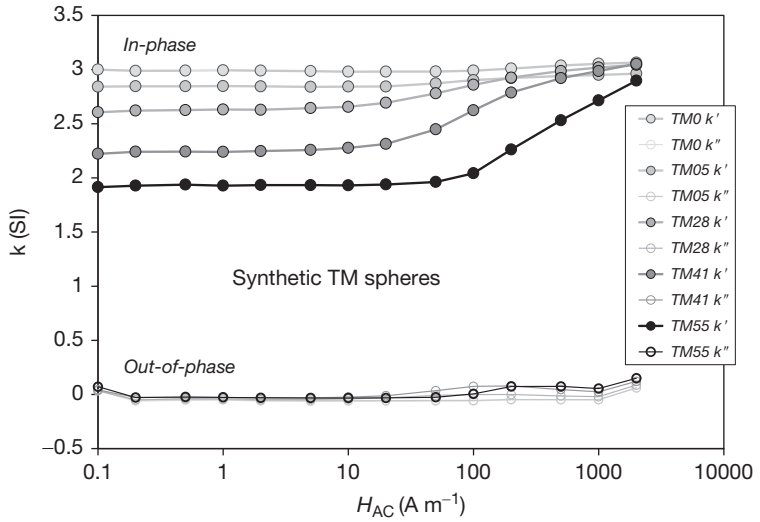


Figure 13 AC field-dependent complex susceptibility in synthetic titanomagnetite spheres of different compositions ($Fe_{3-x}Ti_xO_4$, $0 \leq x \leq 0.55$), due to low-field hysteresis (Jackson et al., 1998). For TM55, the in-phase susceptibility in an AC field of 1000 A m^{-1} exceeds the constant low-field value by almost 50%. Solid (open) symbols indicate in-phase (out-of-phase) susceptibility.

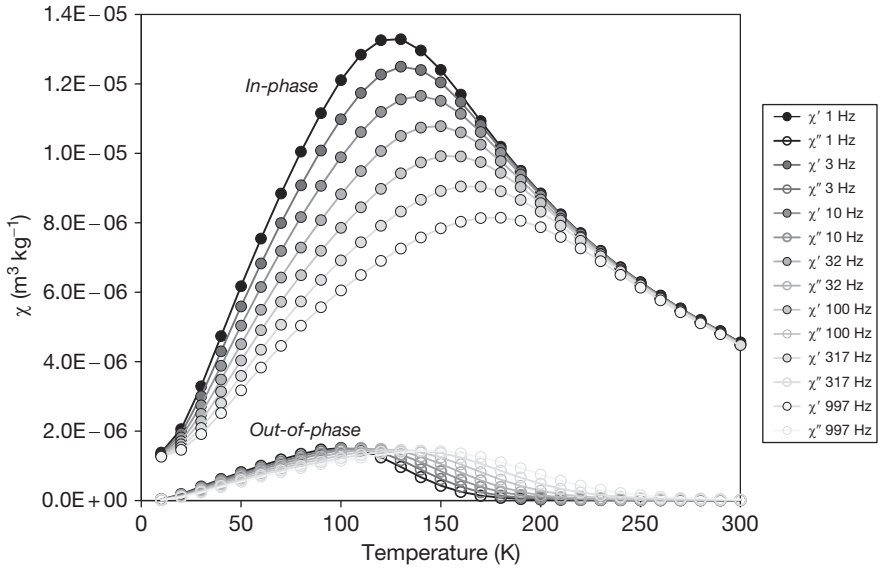


Figure 14 In-phase (χ') and out-of-phase (χ'') susceptibility as functions of AC frequency and temperature for a sample of the Tiva Canyon Tuff, with a narrow size distribution of ferrimagnetic nanoparticles. The ferrimagnetic mineral is a titanium-poor ($x \sim 0.1$) titanomagnetite. Solid (open) symbols indicate in-phase (out-of-phase) susceptibility.

making it more difficult to discriminate between soil and metal anomalies, increasing the number of false-positives, and reducing detection depths. For example, Billings et al (2003) reported that up to 30% of EMI anomalies at a former naval range at Kaho'olawe, Hawaii, are false-positives resulting from magnetic viscosity of ferrimagnetic soil particles at the site. This can be an even more serious detection problem for land mine clearing operations when searching for 'low-metal' land mines found in recent conflict areas (Hannam and Dearing, 2008).

An empirical measurement of the influence of soil magnetic susceptibility on a particular EMI sensor efficiency is provided by the ground reference height (GRH). This parameter is defined as the height of the sensor above the soil surface at which susceptibility effects become significant as the sensor is brought closer to the soil surface from above (**Figure 15**). The larger the GRH value, the greater the influence of the soil susceptibility on sensor performance and the more likely to produce false-positives. GRH values greater than 100 mm are considered severe, while values less than 10 mm are considered neutral (European Committee for Standardization (CEN), 2008). In a study on the susceptibility of tropical soils from various locations, approximately 33% of the samples measured were classified as having a 'severe' or 'very severe' impact on metal detectors (Igel et al., 2009, 2012).

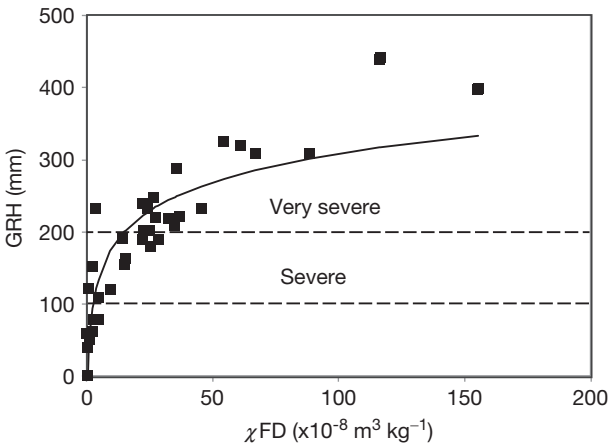


Figure 15 Ground reference height (GRH) as a function of the frequency-dependent susceptibility (xFD) for a Schiebel AN-19/2 metal detector. Field measurements and soil samples taken from the same locations in Croatia, Italy, Laos, and Mozambique. $xFD = x_{LF} - x_{HF}$ is the difference in susceptibility measured at low ($f = 465$ Hz) and high ($f = 4550$ Hz) frequency. Solid line is a log-linear fit to the data ($y = 56.89 \ln(x) + 46.305$, $R^2 = 0.806$). Figure modified from Hannam and Dearing (2008) Mapping soil magnetic properties in Bosnia and Herzegovina for landmine clearance operations. *Earth and Planetary Science Letters* 274: 285–294.

11.05.5 Magnetic Petrology: Induced and Remanent Magnetization of Earth Materials

Magnetic susceptibility and NRM data can provide vital supplements to magnetic investigations of the near-surface. Along with outcrop and drill-hole information, magnetic property data help in relating anomaly signatures to probable sources and provide important constraints for using magnetic anomaly data, both for geologic mapping and for modeling geology at depth. For site-specific investigations, magnetic susceptibility can be used either solely or in conjunction with magnetic or other geophysical data to recognize and investigate subtle anthropogenic features, such as ancient hearths, building foundations, or artificial fill. Although magnetic property measurements taken specifically at the site of investigation are potentially the most useful, it is also instructive to review preexisting databases of magnetic properties, especially at the onset of an investigation. Numerous workers have reviewed magnetic properties of a wide variety of earth materials, including igneous and metamorphic rocks (Clark, 1997, 1999; Clark and Emerson, 1991; Haggerty, 1979; Henkel, 1991), sedimentary rocks (Hansen et al., 2005; Nabighian et al., 2005; Reynolds et al., 1990), soils (Dearing et al., 1996b; Mullins, 1977; Tite and Mullins, 1971), and archaeological substrates (Dalan, 2008). This section will draw upon these and other sources to provide a cursory view of how magnetic properties can contribute to near-surface investigations. For reference, magnetic susceptibilities (**Figure 16**) and Koenigsberger ratios (**Figure 17**) for a variety of rock types are included from a compilation by Clark and Emerson (1991).

11.05.5.1 Igneous Rocks

The magnetic properties of igneous rocks can depend on a broad and somewhat daunting array of factors, including bulk chemistry, initial temperatures of formation, the environment into which the molten rock was emplaced, especially with regard to cooling rate and oxygen fugacity, and the effect of deuteric and other alteration that may have followed emplacement (Clark, 1999; Haggerty, 1979). Thus, it is generally not possible to predict magnetic properties from first-order rock descriptions (Clark, 1999).

With regard to bulk composition, Haggerty (1979) pointed out that oxygen fugacity was related to SiO_2 in such a way that felsic suites (granitic and rhyolitic rocks) are more highly oxidized, relative to mafic suites (basaltic and gabbroic rocks); hence, magnetite and hematite would be typically associated with the former, and ulvöspinel and ilmenite would be typically associated with the latter. Thus, it would appear that felsic igneous rocks would be commonly more magnetic than mafic rocks, but the source of the original magma, as well as the wall rock conditions, can offset this relationship. For example, the previous relationship holds for most magnetic granitoids that were derived from preexisting igneous crust (I-type), but for granitoids derived from metasedimentary crust (S-type), conditions are likely to be more reduced, possibly from interaction with carbon, and thus, they may be more associated with nonmagnetic ilmenite (Clark, 1999). In addition, very high levels of oxidation for granitoids will favor formation of weakly magnetic hematite over that of magnetite (Clark, 1997). On the other hand, titanomagnetite in mafic plutonic rocks tends to exsolve upon cooling, into lamellae of magnetite and either ulvöspinel or ilmenite, commonly increasing susceptibility and, owing to an effective reduction in magnetic grain size, increasing NRM intensity (Clark, 1997; Reynolds et al., 1990). Serpentinization of peridotite results in the formation of MD grains of magnetite, which can also significantly enhance the magnetic susceptibility of ultramafic rocks (Clark, 1999; Oufi et al., 2002; Purucker and Whaler, 2007). Pyrrhotite is a common accessory sulfide in mafic intrusions and ilmenite granites and can be strongly ferrimagnetic in its monoclinic form (Clark and Emerson, 1991).

Volcanic rocks commonly are associated with stronger magnetizations than their intrusive equivalents, owing to high-temperature oxidation and exsolution reactions that favor the formation of SD and PSD (titano)magnetite grains capable of hosting a highly stable and strong remanence ($Q_n > 1$) (**Figure 17**). In general, mafic lavas tend to be more magnetic than intermediate or felsic flows, but the latter two can become comparably magnetic (**Figure 16**) if iron content and oxidations levels are sufficiently high (Clark, 1999; Purucker and Whaler, 2007).

Given the earlier information, it appears that a wide variety of igneous rocks can be associated with a wide range of susceptibilities and NRM values, depending on the physical and chemical conditions in place during the emplacement and subsequent cooling of the rocks. Hence, interpretation of magnetic surveys should include locally derived magnetic properties whenever possible (Purucker and Whaler, 2007). Nonetheless, an extensive magnetic properties database that has been acquired over a large area should at least give a general sense of what might be expected for given rock types. **Figures 16** and **17** summarize 18 years of rock property measurements in Australia by CSIRO, as well as compilations of work published elsewhere. Igneous rock groups that include high-susceptibility (200–10 000 microSI) members

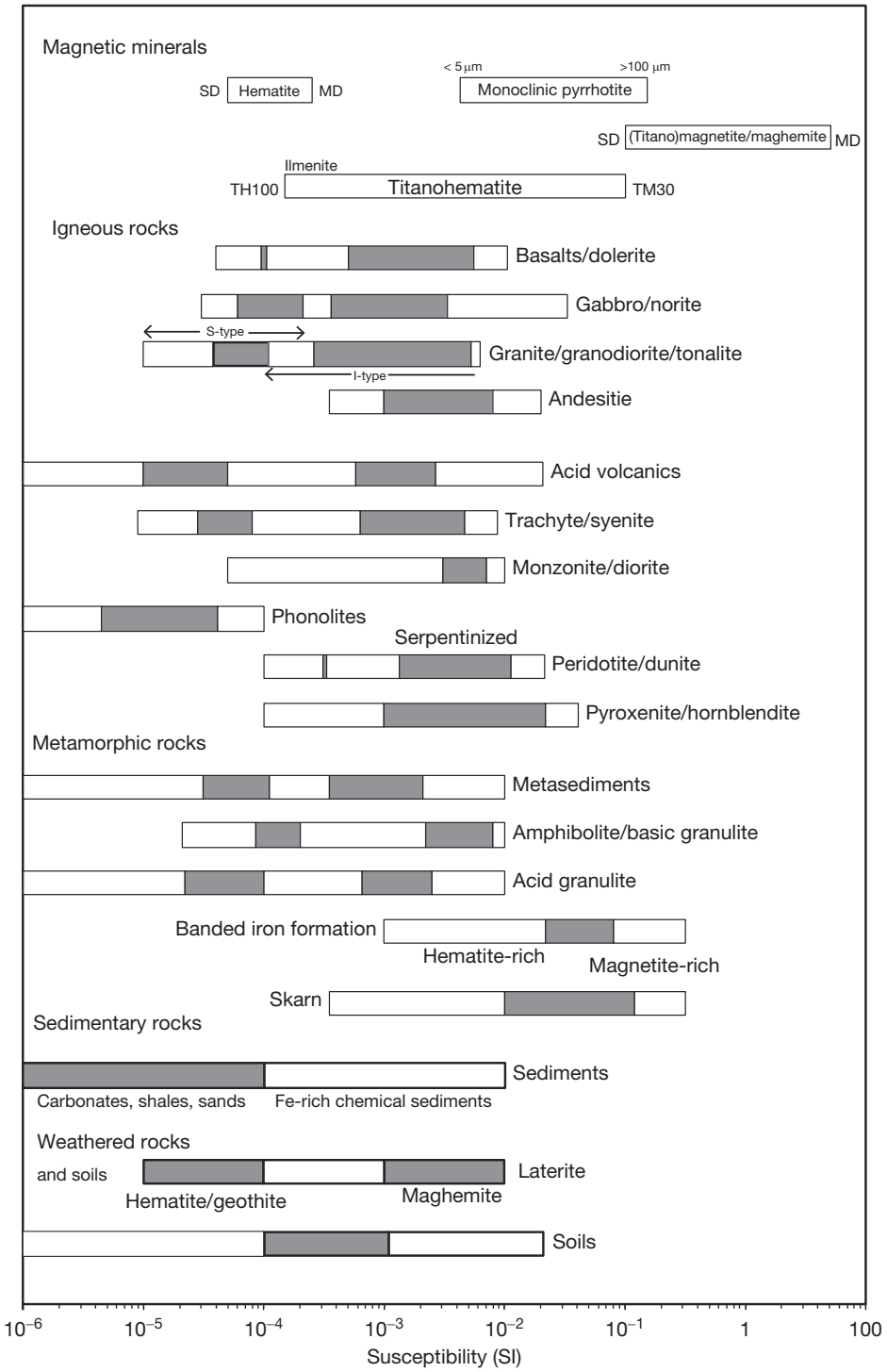


Figure 16 Range of volume magnetic susceptibilities for magnetic minerals and major rock types. Gray portions of bars indicate common susceptibility ranges for various rock types. Note bimodal susceptibility distributions for many rock types. Figure modified and reproduced with permission from the Exploration Geophysics 22(3), 547–555. http://www.publish.csiro.au/nid/224/paper/EG991547.htm (D A Clark and D W Emerson). Copyright Australian Society of Exploration Geophysicists 1991. Published by CSIRO PUBLISHING, Collingwood, Victoria Australia.

are basalt/dolerite, gabbro/norite, granodiorite/tonalite, andesites, acid volcanics, trachyte/syenite, monzonite/diorite, serpentinized peridotite, and pyroxenite/hornblendite. Igneous rock groups that include low magnetic susceptibility members (1–100 microSI) are spilites, some gabbro/norites, ilmenite (S-type) granites, quartz and quartz–feldspar porphyries, trachyte/syenite, and phonolites. Igneous rock groups that include strong NRM values ($Q_n > 1$) are basalt/dolerite (some $Q_n > 100$), gabbro/norite/acid volcanics, andesites/intermediate volcanics, serpentinized peridotite, pyroxenite/hornblendite, and some pyrrhotite-bearing ores (Figure 17). Anorthosites can also be strongly magnetic with high Q_n values

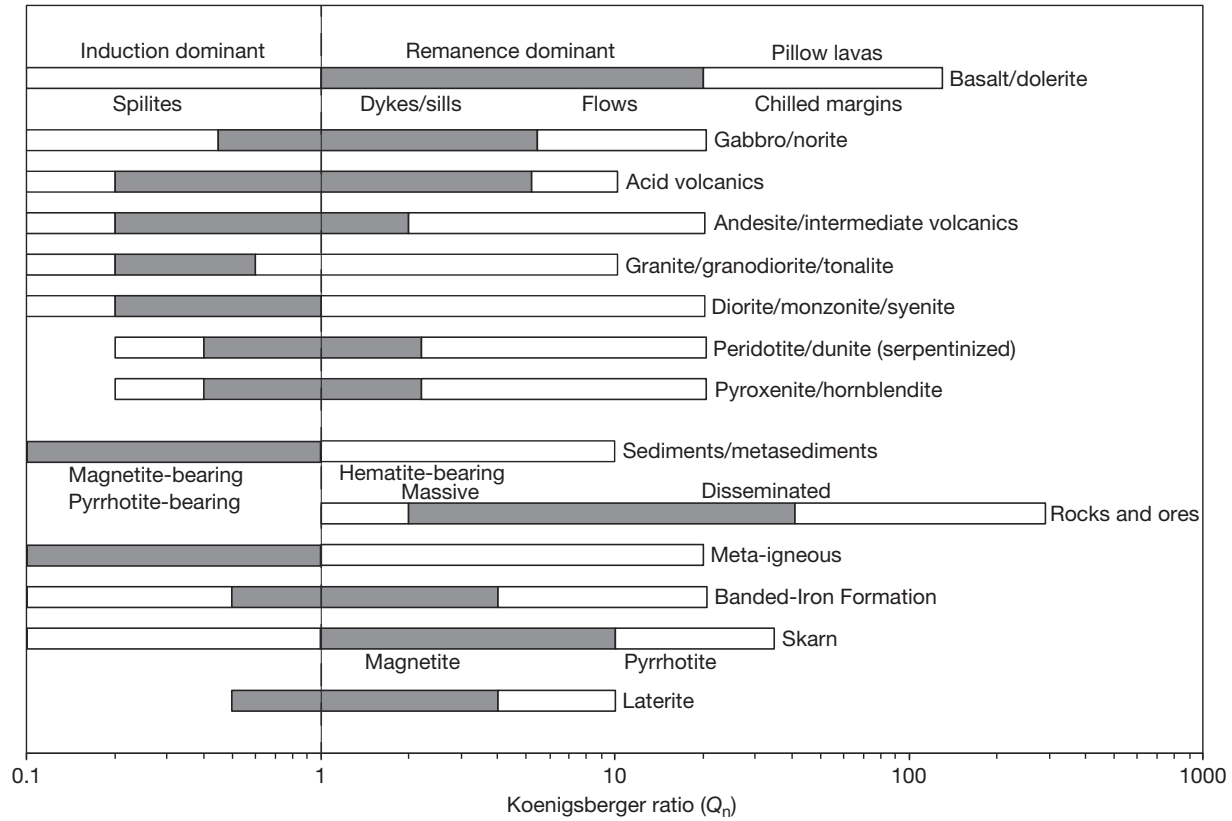


Figure 17 Range of Koenigsberger ratios (Q) for common rock types. Gray portions of bars indicate common ranges. Vertical dashed line designates $Q=1$. Figure modified and reproduced with permission from the Exploration Geophysics 22(3), 547–555. http://www.publish.csiro.au/nid/224/paper/EG991547.htm (D A Clark and D W Emerson). Copyright Australian Society of Exploration Geophysicists 1991. Published by CSIRO PUBLISHING, Collingwood, Victoria Australia.

due to lamellar magnetism in hemoilmenite and, in some cases, result in strong remanent-dominated negative anomalies (Brown and McEnroe, 2004, 2008; McEnroe et al., 2007).

The best-known example of remanence-dominated anomalies ($Q_n > 1$) is the lineated magnetic anomalies in ocean basins associated with seafloor spreading resulting from a combination of oceanic crustal formation at mid-ocean ridges and geomagnetic polarity reversals. These anomalies are produced by a combination of TRM and CRM carried by SD and PSD grains of titanomagnetite ($\sim x = 0.6$) and titanomaghemites in the extrusive layer of pillow basalts of the oceanic crust. A recent review of ocean floor magnetism is given by Gee and Kent (2007).

Many igneous (and metamorphic) groups show a bimodal character, having both low-susceptibility and high-susceptibility members (**Figure 16**). A similar bimodality has been noted in extensive rock property investigations of Precambrian crystalline rocks of the Baltic Shield (Henkel, 1991; Korhonen et al., 2002a,b) and is clearly evident in the density versus magnetic susceptibility summary plot in **Figure 18**. This bimodality is attributed to paramagnetic (low-susceptibility) and ferromagnetic (high-susceptibility) populations that are shared in many rock types (Clark, 1997, 1999; Henkel, 1991). In the paramagnetic population, iron tends to be tied up in silicates instead of oxides, and using the terminology of Henkel (1991), the upper distribution in **Figure 18** defines

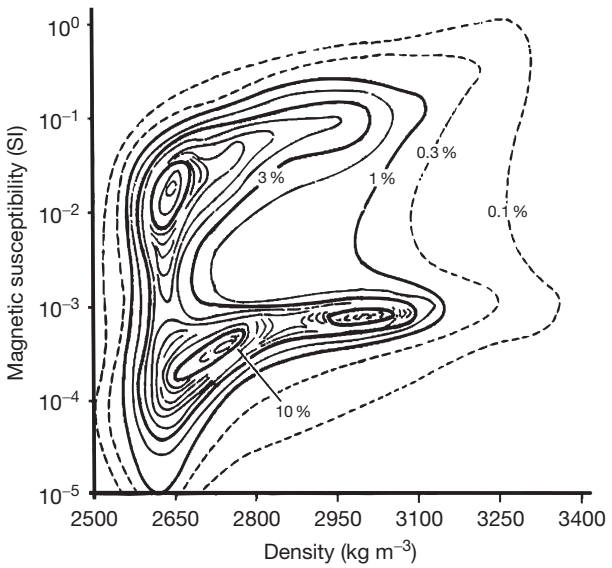


Figure 18 Frequency distribution of 30 000 Precambrian rock samples from the Baltic Shield tested for density and magnetic susceptibility (after Henkel (1991)). Figure modified from Turner GM, Rasson JL, and Reeves CV (2007) Observation and measurement techniques. In: Schubert G (eds.) *Treatise on Geophysics*, Vol. 5. Amsterdam: Elsevier, pp. 93–146.

the 'magnetite trend,' while the lower distribution defines the 'biotite–amphibole paramagnetic trend.' This bimodality may be the cause of the somewhat complex and banded appearance commonly observed over various units in the Baltic Shield (Henkel, 1991) and presumably elsewhere.

11.05.5.2 Sedimentary Rocks

Most carbonate and clastic sedimentary rocks are very weakly magnetic and have been largely ignored by magnetic anomaly interpreters, who were traditionally more interested in estimating either the total thickness of the 'nonmagnetic' sedimentary rocks or the geology of the relatively more magnetic crystalline 'basement' underlying the sedimentary section. However, the rise over the last few decades of higher precision magnetometers and aeromagnetic surveys using denser sampling has revealed subtle magnetic anomalies, generally a few to a few 10 s of nT, arising from the intrasedimentary sources that can reveal significant information on structure, lithology, and possibly even the economic potential of the sedimentary section (Gunn, 1997; Nabighian et al., 2005).

With the exception of iron formations (discussed further later), the susceptibility of most sedimentary rocks are low (1–100 microSI, **Figure 16**) and in many situations are not fully understood (Clark, 1997; Reynolds et al., 1990). Gay (1992) presented examples of intrasedimentary anomalies associated with distributions of detrital magnetites. Reynolds et al. (1990) had pointed out that detrital grains of magnetite and titanomagnetite in sedimentary rocks are reasonably common but highly vulnerable to altering conditions following deposition, such as Eh, pH, temperature, bacterial activity, concentrations of Fe^{2+} and Fe^{3+}, types and abundances of sulfur species, and organic matter. Most of these postdepositional conditions lead to oxidative or reductive replacement of the detrital oxides by essentially nonmagnetic minerals, such as magnetite by hematite or by iron sulfide, respectively. In some sulfide-reducing environments, however, authigenesis of ferrimagnetic sulfide minerals, monoclinic pyrrhotite (Fe_7S_8) and cubic greigite (Fe_3S_4), may occur (Reynolds et al., 1990). Magnetite can be locally produced by microbial activity or destroyed by sulfidization (Goldhaber and Reynolds, 1991) in processes related to hydrocarbon migration, although it is still debated whether this effect can be detected from aeromagnetic surveys (Gay, 1992; Reynolds et al., 1990). In a recent study on the effects of oil biodegradation on the magnetic properties of reservoir host rocks, Emmerton et al. (2013) found an overall decrease in susceptibility with decreasing oil quality resulting from a combination of dissolution of SD magnetite concomitant with the formation of MD magnetite. Thus, the nature of sedimentary-rock magnetism can be complex, and like the study of igneous rocks, it is a good idea to conduct petrographic and rock magnetic studies on some area-specific samples, if available, before interpreting sedimentary-rock-related magnetic anomalies in a new study area.

The great exceptions to the weakly magnetic characterization for most sedimentary rocks are iron formations, which are Archean to Paleoproterozoic precipitates that are generally held to have been deposited in seas that were enriched in ferrous iron, at a time when Earth's atmosphere and hydrosphere were much poorer in free oxygen. Iron formations not only consist chiefly of interlayered iron oxides and silica but also include sulfide, silicate, and carbonate facies (McIntrye, 1980). Oxide-facies iron formations favoring magnetite can be extremely magnetic with susceptibilities generally ranging from 20 000 to 200 000 microSI (**Figure 14**). The transition toward hematite in oxide-facies iron formations, or toward sulfide- and carbonate-facies iron formations, would logically be associated with a significant decrease in magnetic susceptibility, as would the extreme effects of alteration or weathering. Self-demagnetization becomes significant for iron formations with a strong induced component, such that the induced magnetization perpendicular to the layering can be reduced by factors 2–4, relative to magnetization along the layering (Clark and Schmidt, 1994). Iron formations can also possess strong NRMs ($Q_n > 1$) with directions that are significantly different than the induced magnetization making interpretation of magnetic anomalies associated with these formations complicated (Bath, 1962).

11.05.5.3 Metamorphic Rocks

Metamorphism can have a profound effect on magnetic properties of sedimentary and igneous rocks, and a proper magnetic interpretation should take into account metamorphic grade and its possible variation across a study area. The effects of metamorphism on magnetic properties have been discussed at length (Clark, 1997; Grant, 1985a; McIntrye, 1980), and the reader is referred to these references for any information beyond the treatment that is presented here.

Mafic igneous rocks such as basalts usually have sufficient accessory (titano)magnetite so that they are moderately to strongly magnetic when fresh (**Figures 16** and **17**). These magnetite grains tend to remain during zeolite to prehnite–pumpellyite-grade metamorphism, in the absence of hydrothermal fluids. Upon greenschist facies metamorphism, the rock becomes nonmagnetic as the magnetite grains are converted into chlorite, epidote, and hematite (Purucker and Whaler, 2007). In turn, these minerals give way to biotite and amphibole in the amphibolite facies of regional metamorphism. In mafic plutonic rocks such as gabbro, magnetite may survive where it occurs as fine grains within silicate hosts or as an abundant cumulus phase, so that the magnetism of mafic plutonic rocks may be less sensitive to low- and medium-grade metamorphism than that in their extrusive and hypabyssal equivalents (Clark, 1997). Magnetite reappears during granulite-grade metamorphism, and at the highest metamorphic grade (eclogite), the iron returns to silicates such as clinopyroxene and garnet (Purucker and Whaler, 2007). Uplift and decompression of granulites can produce fine-grained magnetite by breakdown of garnet and clinopyroxene, and this magnetite can carry a strong, stable remanence. In general, however, metamorphic magnetite occurs in coarse, multidomain grains that are associated with low Q_n values (**Figure 17**). In contrast, hematite–ilmenite intergrowths in granulite-facies rocks from the Adirondacks (New York, USA), Labrador (Canada), Norway, and Sweden have been shown to have high magnetizations with stable remanence and $Q_n = 10–100$ (Kletetschka and Stout, 1998; Kletetschka et al., 2002; McEnroe and Brown, 2000; McEnroe et al., 2001b).

The source of this stable remanence and associated remanent-dominated anomalies is lamellar magnetism in the nanoscale intergrowths of hematite–ilmenite (McEnroe et al., 2009).

A wide variety of factors affect the magnetization of metamorphosed sedimentary rocks. The iron content of sediments, which is generally higher for pellites than for psammites, and the redox conditions during deposition and diagenesis will largely determine the capacity of sedimentary rocks to develop secondary magnetite during metamorphism (McIntrye, 1980). Significant chemical input by exhalative metal-bearing solutions, including iron, also enhances likelihood for magnetite formation during subsequent metamorphism. Magnetite formation also depends on the oxidation state during metamorphism, which should be intermediate between that which favors ilmenite and that which favors hematite. Magnetite can be produced from high-grade metamorphism of hematite-bearing sediments (Clark, 1997). Finally, metamorphism of organic-rich sediments usually produces graphitic rocks that lack magnetite but may include magnetic pyrrhotite. In the end, magnetic anomaly patterns observed over metasedimentary rocks reflect both sedimentary-facies variations as well as metamorphic conditions (McIntrye, 1980).

11.05.5.4 Unconsolidated Surficial Materials

Unconsolidated surficial materials, which include Pleistocene glacial deposits, alluvium, lacustrine deposits, loess, soils, and paleosols that develop on these materials, were traditionally viewed as the 'overburden' that most geophysical investigations were trying to see through. Consequently, magnetic properties of unconsolidated surficial deposits were usually ignored in most early studies. More recent investigations, however, have revealed that magnetic properties of surficial materials can yield a wealth of information regarding depositional and postdepositional environments, and the last three decades have seen the rise of a powerful subdiscipline of rock magnetism – environmental magnetism, which has made numerous contributions to studies of climate, sedimentology, oceanography, and anthropogenic pollution (Evans and Heller, 2003; Thompson and Oldfield, 1986; Verosub and Roberts, 1995). In addition, the same time period has seen increasing use of rock magnetism methods being successfully applied to archaeological investigations (Dalan and Banerjee, 1996), demining operations (Dabas and Skinner, 1993; Das, 2006; Hannam and Dearing, 2008; Igel et al., 2012), and industrial pollution (Petrovský and Ellwood, 1999; Sagnotti et al., 2006). In short, magnetic properties of once-chided 'overburden' of surficial materials have ended up in many cases providing important information about the composition of these materials and processes responsible for them.

11.05.5.4.1 Magnetic properties of glacial deposits

Pleistocene glacial deposits were typically viewed as magnetically transparent, even though early workers using ground-based magnetic surveys were well aware of anomalies related to glacial deposits (e.g., Heiland, 1946). Recent investigations of glacial deposits have revealed significant variations in magnetic properties, and some of these variations are sufficient to produce detectable anomalies in high-resolution aeromagnetic data (Gay, 2004). Davies et al. (2004) connected channel-like magnetic anomalies to buried paleochannels filled with Tertiary fluvial and Pleistocene glacial deposits; magnetic susceptibilities from the Tertiary and Pleistocene fill ranged from 200 to 800 microSI, while susceptibilities from the enclosing bedrock yielded generally lower values ranging from 50 to 400 microSI. In east-central Minnesota, elevated magnetic susceptibility values ranging from 1250 to 3500 microSI are common for some Pleistocene till and outwash deposits (Chandler et al., 2002). Where the bedrock consists of essentially nonmagnetic sandstone, buried bedrock valleys filled with the relatively magnetic glacial material are easily traceable as sinuous highs (a few to a few 10 s nT) in high-resolution aeromagnetic data. These elevated susceptibilities likely reflect a relative abundance of magnetic rock fragments and oxide grains that were derived from iron formations and mafic igneous rocks, which lie at the bedrock surface directly up-ice. In a separate study of glacial deposits in east-central Minnesota, Chernicoff (1984) demonstrated that magnetic susceptibility could be developed into a useful a tool for discriminating two till units and establishing their differing provenance.

11.05.5.4.2 Magnetic properties of soils and contaminated soils

Soils and paleosols are associated with magnetic property variations that can reflect a variety of natural and anthropogenic conditions (Dearing et al., 1996b; Thompson and Oldfield, 1986; Verosub and Roberts, 1995). The enhancement of magnetic susceptibility values in the upper part of a soil profile, usually horizon A and the upper part of horizon B, has been long known (LeBorgne, 1955), but debate persists as to the causes of secondary ferrimagnetic mineral formation in soils (Blundell et al., 2009; Boyle et al., 2010; Liu et al., 2012; Maher, 2011). Originally, the enhancement was largely attributed to fires, either naturally occurring or anthropogenic, where heating and anaerobic conditions in the upper soil would enhance the formation of magnetic minerals such as magnetite or maghemite from weakly magnetic oxides and hydroxides (Mullins, 1977; Tite and Mullins, 1971). However, it now appears that oxidation/reduction associated with normal pedogenesis through periodic wetting and drying, and perhaps in some cases through bacterial enhancement, may be a more widespread mechanism for producing new magnetic grains in the upper soil (Dearing et al., 1996b). It is now known that the pedogenic grains of magnetite and/or maghemite in the enhanced zone are typically very fine-grained, spanning the SPM–SSD grain size range, and that they may coexist with concentrations of inherited (lithogenic) or anthropogenic magnetic grains, which are coarser (>1 μm) and MD (Maher, 1988; Özdemir and Banerjee, 1982). The high concentrations of nanosized, SPM–SSD particles produced both high susceptibilities and enhanced frequency-dependent susceptibilities.

It has been established that magnetic aerosols produced from different anthropogenic processes including metallurgical production, combustion of fossil fuels, and vehicular traffic contribute to the magnetism of soils in urban areas and downwind from industrial facilities (Hanesch et al., 2007; Hunt et al., 1984; Kapička et al., 2008; Lu and Bai, 2008; Petrovský et al., 2000, 2004; Sagnotti et al., 2006; Spassov et al., 2004; Strzyszcz and Magiera, 1998). The magnetic particles consist

mostly of magnetite, maghemite, and hematite with grain sizes from ~1 to >100 μm (e.g., Petrovský and Ellwood, 1999). Such particles are also usually found in close association with heavy metals (e.g., Cu, Pb, Zn, and Ni) and organic pollutants, and consequently, magnetic methods can serve as quick, low-cost alternatives to geochemical analyses when mapping and assessing soil contamination and map anthropogenic heavy metal pollution from point sources (Durza, 1999; Georgeaud et al., 1997; Hu et al., 2007; Lu and Bai, 2008). A magnetic properties approach may not entirely replace geochemical analysis, but it can at least help direct the chemical follow-up to be more cost-effective (Schmidt et al., 2005).

The most commonly applied magnetic method for contaminated soils is magnetic susceptibility, in part because it can be easily implemented in the field as well as in the lab (Ellwood and Burkart, 1996; Petrovský et al., 2000, 2004; Schmidt et al., 2005). Although considerable variation exists, volume magnetic susceptibilities for contaminated soils typically range from 1000 to >10 000 microSI along with frequency-dependent susceptibilities <3% (Dearing et al., 1996a,b; D'Emilio et al., 2010; Evans and Heller, 2003; Lecoanet et al., 2001). Most results indicate that anthropogenic magnetic particles are typically MD, which distinguishes them from natural pedogenic particles that are finer-grained (<0.1 μm) and within the SPM and SSD size ranges (Dearing et al., 1996a,b; Hu et al., 2007; Lecoanet et al., 2001, 2003). Additional laboratory measurement of coercivity and remanence (typically SIRM and ARM; see next section) is also used to approximate the relative proportion of ferrimagnetic (magnetite and maghemite) and antiferromagnetic (hematite or goethite) components in soil and to identify different contaminating sources of industrial emissions (e.g., Hu et al., 2007; Lecoanet et al., 2003; Petrovský and Ellwood, 1999; Sagnotti et al., 2006). In light of the potential risks posed by contaminated soil, it seems likely that there will be continued use and development of these methods for some time to come.

11.05.5.5 Archaeological Materials

Magnetic methods have long been a part of archaeological investigations, traditionally through ground-based magnetic surveying for site reconnaissance and through archaeomagnetism to help date hearths, fired pottery, and burnt soils. However, recent advances in laboratory-based soil magnetic methods over the last 25 years have added new tools for archaeological investigations (Dalan, 2008; Dalan and Banerjee, 1996; Linford, 2004, 2005). The natural pedogenic processes discussed earlier produce a magnetization background, upon which the further enhancements and degradations are imposed by anthropogenic activities. Of these newer archaeological applications, measurements of susceptibility, frequency-dependent susceptibility (χ_{fd}), and ARM have received much of the attention (Dalan and Banerjee, 1996; Dearing et al., 1996a; Linford, 2004) because these parameters are relatively fast and inexpensive to measure, are minimally destructive, and are sensitive to the concentration and size distribution of magnetic grains in the subsurface can be strongly influenced by anthropogenic activities.

Field measurement of magnetic susceptibility has increased significantly over the last decade, primarily because new instruments and software allow rapid and effective acquisition and interpretation and because magnetic susceptibility has a well-proven record in detecting alterations to soils related to anthropogenic activity (Dalan and Banerjee, 1996; Dearing et al., 1996a; Linford, 2004, 2005). Although intimately related to magnetic anomaly signatures, interpretation of magnetic susceptibility data is inherently simpler than that for magnetic anomaly data. Magnetic anomaly data are directional and dipolar in nature; thus, in inclined fields, the anomaly peaks are shifted relative to their source, and they are associated with flanking minima. In contrast, magnetic susceptibility anomalies approximate those from a monopolar source; the peak is directly over its source area, and no flanking minima occur. In addition, magnetic susceptibility mapping can track transitional variations in magnetic susceptibility, while magnetic anomaly data usually do not. A review of field-based magnetic susceptibility instruments and their applications in the United States is given by Dalan (2008).

A useful magnetic tool for archaeological investigations is an ARM–χ plot (Banerjee et al., 1981; King et al., 1982). Both ARM and χ are sensitive to concentration of magnetic grains, but ARM is particularly sensitive to the presence of finer magnetic grains (SD and PSD), whereas χ is relatively more sensitive to larger grains (larger PSD and MD). This provides a rapid means of characterizing the relative grain sizes and the concentrations of magnetic grains related to natural and archaeological soils. **Figure 19** is an example of an ARM–χ plot for a series of known soil types at Cahokia, which is a major Mississippian (AD 800–1400) cultural site in southern Illinois (Dalan and Banerjee, 1996). Swale soils and ridge soils represent the natural substrate of the site, whereas fill and midden soils represent materials that were artificially emplaced either for filling depressions for a plaza or for building mounds. The major soil types are clearly segregated by this mode of presentation, and the results allowed recognition of soil types, and in some cases mixtures of soil types, in areas where the soil types were

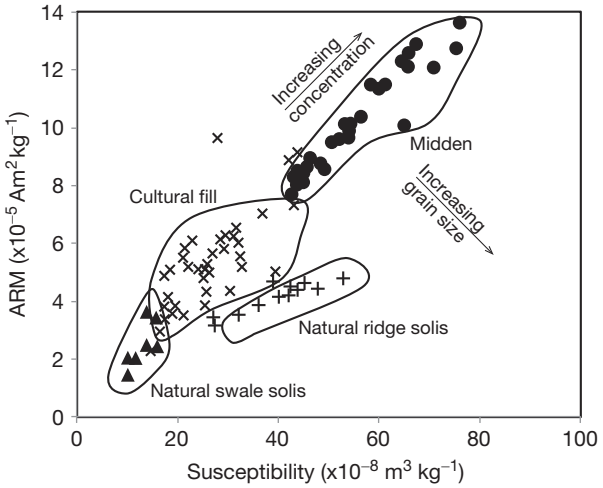


Figure 19 Anhysteretic remanent magnetization (ARM) versus susceptibility for various natural and anthropogenic soils from the Cahokia site. The plot shows relative variations in magnetite concentration and grain size. Figure modified from Dalan and Banerjee (1996) Soil magnetism, an approach for examining archeological landscapes. *Geophysical Research Letters* 23: 185–188.

unknown. The results of the study indicated a dramatically greater amount of moving earth materials had occurred here than previously suspected.

11.05.6 Summary

The magnetic signatures of iron mineral composition and grain size are complex, involving interactions of magnetic domains and walls with crystal defects and microstructures in large grains; vortex and other spin structures in smaller grains; and simple uniform magnetization for only a very narrow range of fine grain sizes. Nanoparticles often have properties and behavior that are fundamentally different from those of the corresponding bulk material. Magnetic properties (including natural remanence) of the iron mineral assemblage in the subsurface represent the end products of its integrated thermomagnetic, biogeochemical, and anthropogenic history. Therefore, magnetic analyses using both field- and laboratory-based methods provide important (though generally indirect) information on the whole set of conditions and processes involved in the formation and subsequent alteration of magnetic materials in the environment and subsurface. Because of the breadth and complexity of atmospheric, biological, chemical, geologic, and anthropogenic processes affecting the characteristics of magnetic minerals (and their memories of past magnetic fields), and because of the richness of the physics that makes these characteristics detectable, the magnetic archive is not only rich but also strongly encrypted. The grand challenges in paleomagnetism, rock magnetism, and environmental magnetism are related to the mining of this archive and deciphering its stored records.

Acknowledgments

The Institute for Rock Magnetism is supported by the NSF Instrumentation and Facilities Program, Earth Science Division. This is IRM contribution 1301. We thank S. McEnroe for providing images for **Figure 5** and two anonymous reviewers for improving this manuscript.

References

Allègre C, Manhès G, and Lewin É (2001) Chemical composition of the Earth and the volatility control on planetary genetics. *Earth and Planetary Science Letters* 185: 49–69.

Aragon R and McCallister RH (1982) Phase and point defect equilibria in the titanomagnetite solid solution. *Physics and Chemistry of Minerals* 8(3): 112–120.

Balsley JR and Buddington AF (1958) Iron titanium oxide minerals, rocks and aeromagnetic anomalies of the Adirondack area. *Economic Geology* 53: 777–805.

Banerjee SK, King JW, and Marvin JA (1981) A rapid method for magnetic granulometry with applications to environmental studies. *Geophysical Research Letters* 8: 333–336.

Bath GD (1962) Magnetic anomalies and magnetizations of the Biwabik Iron-Formation, Mesabi area, Minnesota. *Geophysics* 1: 627–650.

Bazylinski DA and Moskowitz BM (1997) Microbial biomineralization of magnetic iron minerals: Microbiology, magnetism, and environmental significance. In: Banfield JF and Nealson KH (eds.) *Geomicrobiology: Interactions between microbes and minerals*. *Reviews in Mineralogy*, vol. 35, pp. 181–224. Washington, DC: Mineralogical Society of America.

Besnus MJ and Meyer AJ (1964) Nouvelles données expérimentales sur le magnétisme de la pyrrhotine naturelle. In: *International Conference on Magnetism, Nottingham, Institute of Physics and the Physical Society, London*, pp. 507–511.

Bilardello D, Jezek J, and Gilder SA (2013) Role of spherical particles on magnetic field recording in sediments: Experimental and numerical results. *Physics of the Earth and Planetary Interiors* 214: 1–13. http://dx.doi.org/10.1016/j.pepi.2012.10.014.

Bilardello D and Kodama KP (2010) Rock magnetic evidence for inclination shallowing in the early Carboniferous Deer Lake Group red beds of western Newfoundland. *Geophysical Journal International* 181(1): 275–289.

Billings S, Pasion L, Oldenburg D, and Foley J (2003) The influence of magnetic viscosity on electromagnetic sensors. In: *EUDEM-SCOT2, International Conference on Requirements and Technologies for the Detection, Removal and Neutralization of Landmines and UXO* .

Bina M, Tanguy JC, Hoffmann V, et al. (1999) A detailed magnetic and mineralogical study of self-reversed dacitic pumices from the 1991 Pinatubo eruption (Philippines). *Geophysical Journal International* 138(1): 159–178.

Blundell A, Dearing J, Boyle J, and Hannam J (2009) Controlling factors for the spatial variability of soil magnetic susceptibility across England and Wales. *Earth-Science Reviews* 95: 158–188. http://dx.doi.org/10.1016/j.earscirev.2009.05.001.

Borradaile GJ and Jackson MJ (2010) Structural geology, petrofabrics and magnetic fabrics (AMS, AARM, AIRM). *Journal of Structural Geology* 32(10): 1519–1551.

Boyle JF, Dearing JA, Blundell A, and Hannam JA (2010) Testing competing hypotheses for soil magnetic susceptibility using a new chemical kinetic model. *Geology* 38(12): 1059–1062.

Brown LL and McEnroe SA (2004) Paleomagnetism of the Egersund-Ogna anorthosite, Rogaland, Norway, and the position of Fennoscandia in the Late Proterozoic. *Geophysical Journal International* 158(2): 479–488.

Brown LL and McEnroe SA (2008) Magnetic properties of anorthosites: A forgotten source for planetary magnetic anomalies? *Geophysical Research Letters* 35(2).

Burton BP (1991) The interplay of chemical and magnetic ordering. *Reviews in Mineralogy* 25: 303–321.

Burton BP, Robinson P, McEnroe SA, Fabian K, and Boffa Ballaran T (2008) A low-temperature phase diagram for ilmenite-rich compositions in the system Fe_2O_3–$FeTiO_3$. *American Mineralogist* 93: 1260–1272.

Butler RF (1992) *Paleomagnetism: Magnetic Domains to Geologic Terranes*. Oxford: Blackwell.

Chandler VW, Boerboom TJ, and Lively RS (2002) Investigation of stream-like magnetic anomalies in Pine County, Minnesota. In: Boerboom TJ (ed.) *Contributions to the Geology of Pine County Minnesota*. *Report of Investigations*pp. 43–53. Minnesota: Minnesota Geological Survey.

Chang L, Roberts AP, Tang Y, Rainford BD, Muxworthy AR, and Chen QW (2008) Fundamental magnetic parameters from pure synthetic greigite (Fe_3S_4). *Journal of Geophysical Research* 113(B6). http://dx.doi.org/10.1029/2007JB005502.

Charilaou M, Loffler JF, and Gehring AU (2012) Magnetic thermodynamics as proxy for chemical inhomogeneity in hemo-ilmenite solid solutions. *Physics and Chemistry of Minerals* 39(2): 87–92. http://dx.doi.org/10.1007/s00269-011-0463-y.

Chernicoff SE (1984) Using isotropic magnetic-susceptibility to delineate glacial tills. *Journal of Geology* 92(1): 113–118.

Clark DA (1997) Magnetic petrophysics and mineral petrology: Aids to geological interpretation of magnetic surveys. *Journal of Australian Geology and Geophysics* 17(2): 83–103.

Clark DA (1999) Magnetic petrology of igneous intrusions: Implications for exploration and magnetic interpretation. *AGSO Journal of Australian Geology and Geophysics* 17: 83–103.

Clark DA and Emerson DW (1991) Notes on rock magnetization characteristics in applied geophysical studies. *Exploration Geophysics* 22: 547–555.

Clark DA and Schmidt PW (1994) Magnetic properties and magnetic signatures of BIFs of the Hamersley Basin an Yilgarn Block, Western Australia. In: Denith ME, Frankombe KF, Ho SE, Shepherd JM, Groves DI, and Trench A (eds.) *Geophysical Signatures of Western Australia Mineral Deposits*, pp. 343–354. Geology and Geophysics Department (Key Centre) and UWA Extension/The University of Western Australia, Publication No. 26.

Dabas M and Skinner JR (1993) Time-domain magnetization of soils (VRM) experimental relationship to quadrature susceptibility. *Geophysics* 58: 326–333.

Dalan RA (2008) A review of the role of magnetic susceptibility in archaeogeophysical studies in the USA: Recent developments and prospects. *Archaeological Prospection* 15: 1–31.

Dalan RA and Banerjee SK (1996) Soil magnetism, an approach for examining archeological landscapes. *Geophysical Research Letters* 23(2): 185–188.

Das Y (2006) Effects of soil electromagnetic properties on metal detectors. *IEEE Transactions on Geoscience and Remote Sensing* 44: 1444–1453.

Davies J, Mushayandebvu MF, and Smith R (2004) Magnetic detection and characterization of Tertiary and Quaternary buried channels. In: *74th Annual International Meeting, SEG*, 734.

de Boer CB and Dekkers MJ (1996) Grain-size dependence of the rock magnetic properties for a natural maghemite. *Geophysical Research Letters* 23(20): 2815–2818. http://dx.doi.org/10.1029/96gl00781.

Dearing JA, Dann RJL, Hay K, et al. (1996) Frequency-dependent susceptibility measurements of environmental materials. *Geophysical Journal International* 124: 228–240.

Dearing JA, Hay KL, Baban SMJ, Huddleston AS, Wellington EMH, and Loveland PJ (1996) Magnetic susceptibility of soil: An evaluation of conflicting theories using a national data set. *Geophysical Journal International* 127(3): 728–734.

Dekkers MJ and Linssen JH (1989) Rockmagnetic properties of fine-grained natural low temperature haematite with reference to remanence acquisition mechanisms in red beds. *Geophysical Journal International* 99: 1–18.

D'Emilio M, Caggiano R, Coppola R, Macchiato M, and Ragosta M (2010) Magnetic susceptibility measurements as proxy method to monitor soil pollution: The case study of S. Nicola di Melfi. *Environmental Monitoring and Assessment* 169: 619–630.

Denham CR and Chave AD (1982) Detrital remanent magnetization: Viscosity theory of the lock-in zone. *Journal of Geophysical Research B: Solid Earth* 87: 7126–7130.

Dunlop DJ (1983) Viscous magnetization of 0.04–100 μm magnetites. *Geophysical Journal of the Royal Astronomical Society* 74: 667–687.

Dunlop DJ (1986) Hysteresis properties of magnetite and their dependence of particle size: A test of pseudo-single-domain remanence models. *Journal of Geophysical Research* 91(B9): 9569–9584.

Dunlop DJ and Arkani-Hamed J (2005) Magnetic minerals in the Martian crust. *Journal of Geophysical Research: Planets* 110(E12): 1–11. http://dx.doi.org/10.1029/2005je002404.

Dunlop DJ and Kletetschka G (2001) Multidomain hematite: A source of planetary magnetic anomalies? *Geophysical Research Letters* 28(17): 3345–3348.

Dunlop DJ and Özdemir Ö (1997) *Rock Magnetism: Fundamentals and Frontiers*. Cambridge: Cambridge University Press.

Dunlop DJ and Özdemir Ö (2007) Magnetization in rocks and minerals. In: Schubert G (ed.) *Geomagnetism: Treatise on Geophysics*, vol. 5, pp. 277–336. Amsterdam: Elsevier.

Durza O (1999) Heavy metals contamination and magnetic susceptibility in soils around metallurgical plan. *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy* 24(6): 541–543.

Egli R (2004a) Characterization of individual rock magnetic components by analysis of remanence curves. 1. Unmixing natural sediments. *Studia Geophysica et Geodaetica* 48: 391–446.

Egli R (2004b) Characterization of individual rock magnetic components by analysis of remanence curves. 2. Fundamental properties of coercivity distributions. *Physics and Chemistry of the Earth* 29(13–14): 851–867.

Egli R (2004c) Characterization of individual rock magnetic components by analysis of remanence curves. 3. Bacterial magnetite and natural processes in lakes. *Physics and Chemistry of the Earth* 29(13–14): 869–884.

Egli R (2006) Theoretical considerations on the anhysteretic remanent magnetization of interacting particles with uniaxial anisotropy. *Journal of Geophysical Research B: Solid Earth* 111: B12S18. http://dx.doi.org/10.1029/2006JB004577.

Ellwood BB and Burkart B (1996) A test of hydrocarbon-induced magnetic patterns in soils: The sanitary landfill as laboratory. In: Schumacher D and Abrams MA (eds.) *Hydrocarbon Migration and Its Near-Surface Expression*, *AAPG Memoir* 66: 91–98.

Emmerton S, Muxworthy AR, Sephton MA, et al. (2013) Correlating biodegradation to magnetization in oil bearing sedimentary rocks. *Geochimica et Cosmochimica Acta* 112: 146–165.

Eppelbaum LV (2011) Study of magnetic anomalies over archaeological targets in urban environments. *Physics and Chemistry of the Earth* 36(16): 1318–1330. http://dx.doi.org/10.1016/j.pce.2011.02.005.

Essene EJ and Fisher DC (1986) Lightning strike fusion – Extreme reduction and metal-silicate liquid immiscibility. *Science* 234(4773): 189–193. http://dx.doi.org/10.1126/science.234.4773.189.

Evans ME and Heller F (2003) *Environmental Magnetism: Principles and Applications of Environmagnetics*. Amsterdam: Academic Press, Elsevier Science.

Fabian K (2009) Thermochemical remanence acquisition in single-domain particle ensembles: A case for possible overestimation of the geomagnetic paleointensity. *Geochemistry, Geophysics, Geosystems* 10: Q06Z03. http://dx.doi.org/10.1029/2009GC002420.

Fabian K, Kirchner A, Williams W, Heider F, Leibl T, and Huber A (1996) Three-dimensional micromagnetic calculations for magnetite using FFT. *Geophysical Journal International* 124: 89–104.

Fabian K, McEnroe SA, Robinson P, and Shcherbakov VP (2008) Exchange bias identifies lamellar magnetism as the origin of the natural remanent magnetization in titanohematite with ilmenite exsolution from Modum, Norway. *Earth and Planetary Science Letters* 268: 339–353.

Fuller M, Kidane T, and Ali J (2002) AF demagnetization characteristics of NRM, compared with anhysteretic and saturation isothermal remanence: An aid in the interpretation of NRM. *Physics and Chemistry of the Earth* 27(25–31): 1169–1177.

Gay SP Jr. (1992) Epigenetic versus syngenetic magnetite as a cause of magnetic anomalies. *Geophysics* 57: 60–68.

Gay SP Jr. (2004) Glacial till: A troublesome source of near-surface magnetic anomalies. *The Leading Edge* 23: 542–547.

Gee JS and Kent DV (2007) Source of oceanic magnetic anomalies and the geomagnetic polarity timescale. In: Schubert G (ed.) *Geomagnetism: Treatise on Geophysics*, vol. 5, pp. 455–507. Amsterdam: Elsevier.

Gehring AU, Fischer H, Louvel M, Kunze K, and Weidler PG (2009) High temperature stability of natural maghemite: A magnetic and spectroscopic study. *Geophysical Journal International* 179(3): 1361–1371. http://dx.doi.org/10.1111/j.1365-246X.2009.04348.x.

Gendler TS, Shcherbakov VP, Dekkers MJ, Gapeev AK, Gribov SK, and McClelland E (2005) The lepidocrocite–maghemite–haematite reaction chain-I. Acquisition of chemical remanent magnetization by maghemite, its magnetic properties and thermal stability. *Geophysical Journal International* 160(3): 815–832.

Georgeaud VM, Rochette P, Ambrosi JP, Vandamme D, and Williamson D (1997) Relationship between heavy metals and magnetic properties in a large polluted catchment: The Etang de Berre (South of France). *Physics and Chemistry of the Earth* 22(1–2): 211–214.

Gilder S (2007) Magnetization, piezoremanence and stress demagnetization. In: Gubbins D and Herrero-Bervera E (eds.) *Encyclopedia of Geomagnetism and Paleomagnetism*, pp. 599–603. Netherlands: Springer.

Goldhaber M and Reynolds RL (1991) Relations among hydrocarbon reservoirs, epigenetic sulfidization, and rock magnetization: Examples from the south Texas coastal plain. *Geophysics* 56: 748–757.

Grant FS (1985a) Aeromagnetics, geology, and ore environments, I. Magnetite in igneous, sedimentary, and metamorphic rocks: An overview. *Geoexploration* 23: 303–333.

Grant FS (1985b) Aeromagnetics, geology, and ore environments, II. Magnetite and ore environments. *Geoexploration* 23: 335–362.

Grauch VJS and Hudson MR (2011) Aeromagnetic anomalies over faulted strata. *The Leading Edge* 30: 1242–1252.

Gubbins D and Herrero-Bervera E (eds.) (2007) *Encyclopedia of Geomagnetism and Paleomagnetism*. Netherlands: Springer.

Gunn PJ (1997) Application of aeromagnetic surveys to sedimentary basin studies. *AGSO Journal of Australian Geology and Geophysics* 17: 133–144.

Haggerty SE (1979) The aeromagnetic minerology of igneous rocks. *Canadian Journal of Earth Sciences* 16: 1281–1293.

Haggerty SE and Toft PB (1985) Native iron in the continental lower crust: Petrological and geophysical implications. *Science* 229: 647–649.

Halgedahl S (2007) Magnetic domains. In: Gubbins D and Herrero-Bervera E (eds.) *Encyclopedia of Geomagnetism and Paleomagnetism*, pp. 874–883. Netherlands: Springer.

Hanesch M, Rantitsch G, Hemetsberger S, and Scholger R (2007) Lithological and pedological influences on the magnetic susceptibility of soil: Their consideration in magnetic pollution mapping. *Science of the Total Environment* 382(2–3): 351–363.

Hannam JA and Dearing JA (2008) Mapping soil magnetic properties in Bosnia and Herzegovina for landmine clearance operations. *Earth and Planetary Science Letters* 274(3–4): 285–294. http://dx.doi.org/10.1016/j.epsl.2008.05.006.

Hansen RO, Racic L, and Grauch VJS (2005) Magnetic methods in near-surface geophysics. In: Butler DK (ed.) *Near-Surface Geophysics*, pp. 151–178. Tulsa, OK: Society of Exploration Geophysicists.

Harrison RJ (2006) Microstructure and magnetism in the ilmenite–hematite solid solution: A Monte Carlo simulation study. *American Mineralogist* 91(7): 1006–1023.

Harrison RJ, Dunin-Borkowski RE, Kasama T, Simpson ET, and Feinberg JM (2007) In: Schubert G (ed.) *Magnetic Properties of Rocks and Minerals. Treatise on Geophysics. Vol. 2: Mineral Physics*, pp. 579–630. Amsterdam: Elsevier.

Harrison RJ and Feinberg JM (2009) Mineral magnetism: Providing new insights into geoscience processes. *Elements* 5(4): 209–214.

Harrison RJ and Putnis A (1999) The magnetic properties and crystal chemistry of oxide spinel solid solutions. *Surveys in Geophysics* 19(6): 461–520. http://dx.doi.org/10.1023/a:1006535023784.

Heider F and Dunlop DJ (1987) Two types of chemical remanent magnetization during the oxidation of magnetite. *Physics of the Earth and Planetary Interiors* 46: 24–45.

Heiland CA (1946) *Geophysical Exploration*. New York: Prentice-Hall.

Henkel H (1991) Petrophysical properties (density and magnetization) of rocks from the northern part of the Baltic Shield. *Tectonophysics* 192: 1–19.

Heslop D, Dekkers MJ, Kruiver PP, and van Oorschot IHM (2002) Analysis of isothermal remanent magnetization acquisition curves using the expectation–maximization algorithm. *Geophysical Journal International* 148(1): 58–64.

Heslop D and Dillon M (2007) Unmixing magnetic remanence curves without a priori knowledge. *Geophysical Journal International* 170(2): 556–566.

Heslop D and Roberts AP (2012) A method for unmixing magnetic hysteresis loops. *Journal of Geophysical Research: Solid Earth* 117: B03103. http://dx.doi.org/10.1029/2011jb008859.

Hinze WJ, von Frese RRB, and Saad AH (2013) *Gravity and Magnetic Exploration*. Cambridge: Cambridge University Press.

Hoffman KA (1992) Self-reversal of thermoremanent magnetization in the ilmenite–hematite system: Order–disorder, symmetry, and spin alignment. *Journal of Geophysical Research B: Solid Earth* 97: 10883–10895.

Hofmann DI and Fabian K (2009) Correcting relative paleointensity records for variations in sediment composition: Results from a South Atlantic stratigraphic network. *Earth and Planetary Science Letters* 284(1–2): 34–43.

Hrouda F (2002) Low-field variation of magnetic susceptibility and its effect on the anisotropy of magnetic susceptibility of rocks. *Geophysical Journal International* 150(3): 715–723.

Hu XF, Su Y, Ye R, Li XQ, and Zhang GL (2007) Magnetic properties of the urban soils in Shanghai and their environmental implications. *Catena* 70: 428–436.

Hunt A, Jones J, and Oldfield F (1984) Magnetic measurement and heavy metals in atmospheric particulates of anthropogenic origin. *Science of the Total Environment* 33: 129–139.

Hunt CP, Moskowitz BM, and Banerjee SK (1995) Magnetic properties of rocks and minerals. In: Ahrens TJ (ed.) *A Handbook of Physical Constants*, vol. 3, pp. 189–204. Washington, DC: American Geophysical Union.

Igel J, Preetz H, and Altfelder S (2009) Predicting soil influence on the performance of metal detectors: Magnetic properties of tropical soils. *Journal of Mine Action* (13.1): 103–107.

Igel J, Preetz H, and Altfelder S (2012) Magnetic viscosity of tropical soils: Classification and prediction as an aid for landmine detection. *Geophysical Journal International* 190(2): 843–855.

Ishikawa Y and Syono Y (1963) Order–disorder transformation and reverse thermoremanent magnetism in the $FeTiO_3$–Fe_2O_3 system. *Journal of Physics and Chemistry of Solids* 24: 517–528.

Jackson M, Moskowitz B, Rosenbaum J, and Kissel C (1998) Field-dependence of AC susceptibility in titanomagnetites. *Earth and Planetary Science Letters* 157(3–4): 129–139.

Jackson M and Sølheid P (2010) On the quantitative analysis and evaluation of magnetic hysteresis data. *Geochemistry, Geophysics, Geosystems* 11: Q04Z15. http://dx.doi.org/10.1029/2009GC002932.

Kapička A, Petrovský E, Fialová H, Podrázský V, and Dvořák I (2008) High resolution mapping of anthropogenic pollution in the Giant Mountains National Park using soil magnetometry. *Studia Geophysica et Geodaetica* 52: 271–284.

Kasama T, Harrison RJ, Church NS, Nagao M, Feinberg JM, and Dunin-Borkowski RE (2012) Ferrimagnetic/ferroelastic domain interactions in magnetite below the Verwey transition. Part 1: Electron holography and Lorentz microscopy. *Phase Transitions: A Multinational Journal* 86(1): 67–87. http://dx.doi.org/10.1080/01411594.2012.695373.

Kelso PR and Banerjee SK (1994) Elevated temperature viscous remanent magnetization of natural and synthetic multidomain magnetite. *Earth and Planetary Science Letters* 122: 43–56.

Kelso PR, Banerjee SK, and Worm H-U (1991) The effect of low-temperature hydrothermal alteration on the remanent magnetization of synthetic titanomagnetites: A case for the acquisition of chemical remanent magnetization. *Journal of Geophysical Research B: Solid Earth* 96: 19545–19553.

Kind J, García-Rubio I, Charilaou M, Nowaczyk NR, Löffler JF, and Gehring AU (2013) Domain-wall dynamics in 4C pyrrhotite at low temperature. *Geophysical Journal International* 195(1): 192–199. http://dx.doi.org/10.1093/gji/ggt262.

King JW, Banerjee SK, Marvin JA, and Özdemir Ö (1982) A comparison of different magnetic methods for determining the relative grain size of magnetite in natural materials: Some results from lake sediments. *Earth and Planetary Science Letters* 59: 404–419.

Kittel C (1949) Physical theory of ferromagentic domains. *Reviews of Modern Physics* 21: 541–583.

Klein F and Bach W (2009) Fe–Ni–Co–O–S phase relations in peridotite–seawater interactions. *Journal of Petrology* 50(1): 37–59. http://dx.doi.org/10.1093/petrology/egn071.

Kletetschka G and Stout JH (1998) The origin of magnetic anomalies in lower crustal rocks, Labrador. *Geophysical Research Letters* 25(2): 199–202.

Kletetschka G, Wasilewski PJ, and Taylor PT (2000a) Hematite vs. magnetite as the signature for planetary magnetic anomalies? *Physics of the Earth and Planetary Interiors* 119(3): 259–267.

Kletetschka G, Wasilewski PJ, and Taylor PT (2000b) Unique thermoremanent magnetization of multidomain sized hematite: Implications for magnetic anomalies. *Earth and Planetary Science Letters* 176(3): 469–479.

Kletetschka G, Wasilewski P, and Taylor P (2002) The role of hematite–ilmenite solid solution in the production of magnetic anomalies in ground- and satellite-based data. *Tectonophysics* 347: 167–177.

Korhonen JV, Aaro S, All T, et al. (2002a) *Bouguer Anomaly Map of the Fennoscandian Shield 1:2000000*. Geological Surveys of Finland, Norway and Sweden and Ministry of Natural Resources of Russian Federation.

Korhonen JV, Aaro S, All T, et al. (2002b) *Magnetic Anomaly Map of the Fennoscandian Shield 1:2000000*. Geological Surveys of Finland, Norway and Sweden and Ministry of Natural Resources of Russian Federation.

Krása D and Fabian K (2007) Rock magnetism, hysteresis measurements. In: Gubbins D and Herrero-Bervera E (eds.) *Encyclopedia of Geomagnetism and Paleomagnetism*, pp. 874–883. Netherlands: Springer.

Lagroix F, Banerjee SK, and Moskowitz BM (2004) Revisiting the mechanism of reversed thermoremanent magnetization based on observations from synthetic ferrian ilmenite ($y = 0.7$). *Journal of Geophysical Research* 109(B12). http://dx.doi.org/10.1029/B01104.

Langel RA and Hinze WJ (1998) *The Magnetic Field of the Earth's Lithosphere: The Satellite Perspective*. New York: Cambridge University Press.

Lattard D, Engelmann R, Kontny A, and Sauerzapf U (2006) Curie temperatures of synthetic titanomagnetites in the Fe–Ti–O system: Effects of composition, crystal chemistry, and thermomagnetic methods. *Journal of Geophysical Research: Solid Earth* 111(B12S28). http://dx.doi.org/10.1029/2006JB004591.

LeBorgne E (1955) Susceptibilité magnétique anormale du sol superficiel. *Annales de Geophysique* 11: 399–419.

Lecoanet H, Lévêque F, and Ambrosi J (2001) Magnetic properties of salt-marsh soils contaminated by iron industry emissions in southeast France. *Journal of Applied Geophysics* 48: 67–81.

Lecoanet H, Leveque F, and Ambrosi JP (2003) Combination of magnetic parameters: An efficient way to discriminate soil-contamination sources (south France). *Environmental Pollution* 122: 229–234.

Lindsley DH (ed.) (1991) *Oxide Minerals: Petrologic and Magnetic Significance*. Blacksburg, VA: Mineralogical Society of America.

Linford NT (2004) Magnetic ghosts: Mineral magnetic measurements on Roman and Anglo-Saxon graves. *Archaeological Prospection* 11(3): 167–180.

Linford N (2005) Archaeological applications of naturally occurring nanomagnets. *Journal of Physics: Conference Series* 17(1): 127–144.

Liu QS, Roberts AP, Larrasoana JC, et al. (2012) Environmental magnetism: Principles and applications. *Reviews of Geophysics* 50: Available: WOS:000310960500001, [accessed November].

Liu QS, Roberts AP, Rohling EJ, Zhu RX, and Sun YB (2008) Post-depositional remanent magnetization lock-in and the location of the Matuyama–Brunhes geomagnetic reversal boundary in marine and Chinese loess sequences. *Earth and Planetary Science Letters* 275(1–2): 102–110.

Lowrie W (2007) *Fundamentals of Geophysics*. New York: Cambridge University Press.

Lu SG and Bai SQ (2008) Magnetic characterization and magnetic mineralogy of the Hangzhou urban soils and its environmental implications. *Chinese Journal of Geophysics* 51: 549–557.

Maher BA (1988) Magnetic properties of some synthetic submicron magnetites. *Geophysical Journal* 94: 83–96.

Maher BA (2011) The magnetic properties of Quaternary aeolian dusts and sediments, and their palaeoclimatic significance. *Aeolian Research* 3(2): 87–144. http://dx.doi.org/10.1016/j.aeolia.2011.01.005.

Maher BA, Karloukovski VV, and Mutch TJ (2004) High-field remanence properties of synthetic and natural submicrometre haematites and goethites: Significance for environmental contexts. *Earth and Planetary Science Letters* 226(3–4): 491–505.

Maki D (2005) Lightning strikes and prehistoric ovens: Determining the source of magnetic anomalies using techniques of environmental magnetism. *Geoarchaeology – An International Journal* 20(5): 449–459.

McClelland E (1996) Theory of CRM acquired by grain growth, and its implications for TRM discrimination and paleointensity determination in igneous rocks. *Geophysical Journal International* 126(1): 271–280.

McEnroe SA and Brown LL (2000) A closer look at remanence-dominated aeromagnetic anomalies: Rock magnetic properties and magnetic mineralogy of the Russell Belt microcline–sillimanite gneiss, northwest Adirondack Mountains, New York. *Journal of Geophysical Research* 105(B7): 16437–16456.

McEnroe SA, Brown LL, and Robinson P (2004) Earth analog for Martian magnetic anomalies: Remanence properties of hemo-ilmenite norites in the Bjerkreim-Sokndal intrusion, Rogaland, Norway. *Journal of Applied Geophysics* 56(3): 195–212.

McEnroe SA, Fabian K, Robinson P, Gaina C, and Brown LL (2009) Crustal magnetism, lamellar magnetism and rocks that remember. *Elements* 5(4): 241–246.

McEnroe SA, Harrison RJ, Robinson P, Golla U, and Jercinovic MJ (2001) The effect of fine-scale microstructures in titanohematite on the acquisition and stability of

natural remanent magnetization in granulite facies metamorphic rocks, southwest Sweden: Implications for crustal magnetism. *Journal of Geophysical Research* 106(B12): 30523–30546.

McEnroe SA, Robinson P, Langenhorst F, Frandsen C, Terry MP, and Ballaran TB (2007) Magnetization of exsolution intergrowths of hematite and ilmenite: Mineral chemistry, phase relations, and magnetic properties of hemo-ilmenite ores with micron- to nanometer-scale lamellae from Allard Lake, Quebec. *Journal of Geophysical Research* 112(B10103). http://dx.doi.org/10.1029/2007JB004973.

McEnroe SA, Robinson P, and Panish PT (2001) Aeromagnetic anomalies, magnetic petrology, and rock magnetism of hemo-ilmenite- and magnetite-rich cumulate rocks from the Sokndal Region, South Rogaland, Norway. *American Mineralogist* 86: 1447–1468.

McIntrye JI (1980) Geological significance of magnetic patterns related to magnetite in sediments and metasediments – A review. *Bulletin of the Australian Society of Exploration Geophysics* 11: 19–33.

Merrill RT and McFadden PL (1995) Dynamo theory and paleomagnetism. *Journal of Geophysical Research B: Solid Earth* 100: 317–326.

Mitra R and Tauxe L (2009) Full vector model for magnetization in sediments. *Earth and Planetary Science Letters* 286(3–4): 535–545.

Morin FJ (1950) Magnetic susceptibility of α-Fe_2O_3 and α-Fe_2O_3 with added titanium. *Physical Review B* 78: 819–820.

Moskowitz BM (2007) Anhysteretic remanent magnetization. In: Gubbins D and Herrero-Bervera E (eds.) *Encyclopedia of Geomagnetism and Paleomagnetism*, pp. 572–580. Springer.

Moskowitz BM, Frankel RB, Walton SA, et al. (1997) Determination of the pre-exponential frequency factor for superparamagnetic maghemite particles in magnetoferritin. *Journal of Geophysical Research B: Solid Earth* 102(B10): 22671–22680.

Moskowitz BM, Jackson M, and Kissel C (1998) Low-temperature magnetic behavior of titanomagnetites. *Earth and Planetary Science Letters* 157(3–4): 141–149.

Mullins CE (1977) Magnetic susceptibility of the soil and its significance in soil science – A review. *Journal of Soil Science* 28: 223–246.

Muxworthy AR, Dunlop DJ, and Williams W (2003) High-temperature magnetic stability of small magnetite particles. *Journal of Geophysical Research* 108(B5): EPM 18-1–EPM 18-9.

Nabighian MN, Grauch VJS, Hansen RO, et al. (2005) The historical development of the magnetic method in exploration. *Geophysics* 70(6): 33ND–61ND. http://dx.doi.org/10.1190/1.2133784.

Néel L (1949) Théorie du traînage magnétique des ferromagnétiques en grains fins avec applications aux terres cuites. *Annales de Geophysique* 5: 99–136.

Néel L (1955) Some theoretical aspects of rock magnetism. *Advances in Physics* 4: 191–243.

Nininger RC Jr. and Schroeer D (1978) Mössbauer studies of the Morin transition in bulk and microcrystalline α-Fe_2O_3. *Journal of Physics and Chemistry of Solids* 39: 137–144.

Nord GL Jr. and Lawson CL (1989) Order–disorder transition-induced twin domains and magnetic properties in ilmenite–hematite. *American Mineralogist* 74: 160–176.

Nord GL Jr. and Lawson CA (1992) Magnetic properties of $ilmenite_{70}$–$hematite_{30}$: Effect of transformation-induced twin boundaries. *Journal of Geophysical Research B: Solid Earth* 97: 10897–10910.

O'Reilly W (1984) *Rock and Mineral Magnetism*. Glasgow: Blackie.

Oufi O, Cannat M, and Horen H (2002) Magnetic properties of variably serpentinized abyssal peridotite. *Journal of Geophysical Research* 107(B5). http://dx.doi.org/10.1029/2001JB000549.

Özdemir Ö and Banerjee SK (1982) A preliminary magnetic study of soil samples from west-central Minnesota. *Earth and Planetary Science Letters* 59: 393–403.

Özdemir Ö and Banerjee SK (1984) High temperature stability of maghemite (γ-Fe_2O_3). *Geophysical Research Letters* 11: 161–164.

Özdemir Ö and Dunlop DJ (1985) An experimental study of chemical remanent magnetizations of synthetic monodomain titanomaghemites with initial thermoremanent magnetizations. *Journal of Geophysical Research B: Solid Earth* 90: 11513–11523.

Özdemir Ö and Dunlop DJ (1993) Chemical remanent magnetization during g-FeOOH phase transformations. *Journal of Geophysical Research B: Solid Earth* 98: 4191–4198.

Özdemir Ö and Dunlop DJ (1996) Thermoremanence and Néel temperature of goethite. *Geophysical Research Letters* 23(9): 921–924.

Özdemir Ö and Dunlop DJ (2006) Magnetic memory and coupling between spin-canted and defect magnetism in hematite. *Journal of Geophysical Research B: Solid Earth* 111(B12S03). http://dx.doi.org/10.1029/2006JB004555.

Özdemir Ö, Dunlop DJ, and Moskowitz BM (1993) The effect of oxidation of the Verwey transition in magnetite. *Geophysical Research Letters* 20: 1671–1674.

Petrovský E and Ellwood BB (1999) Magnetic monitoring of air-, land- and water pollution. In: Maher BA and Thompson R (eds.) *Quaternary Climates, Environments and Magnetism*, pp. 279–322. Cambridge: Cambridge University Press.

Petrovský E, Hulka Z, and Kapička A (2004) A new tool for in situ measurements of the vertical distribution of magnetic susceptibility in soils as basis for mapping deposited dust. *Environmental Technology* 25(9): 1021–1029.

Petrovský E, Kapička A, Jordanova N, Knab M, and Hoffmann V (2000) Low-field magnetic susceptibility: A proxy method of estimating increased pollution of different environmental systems. *Environmental Geology* 39: 312–318.

Piekarz P, Parlinski K, and Olés AM (2007) Origin of the Verwey transition in magnetite: Group theory, electronic structure, and lattice dynamics study. *Physical Review B: Condensed Matter and Materials Physics* 76(16). http://dx.doi.org/10.1103/PhysRevB.76.165124.

Pokhil TG and Moskowitz BM (1997) Magnetic domains and domain walls in pseudo-single-domain magnetite studied with magnetic force microscopy. *Journal of Geophysical Research B: Solid Earth* 102(B10): 22681–22694.

Prévot M, Hoffman KA, Goguitchaichvili A, Doukhan J-C, Shcherbakov V, and Bina M (2001) The mechanism of self-reversal of thermoremanence in natural hemoilmenite crystals: New experimental data and model. *Physics of the Earth and Planetary Interiors* 126(1–2): 75–92.

Pullaiah G, Irving E, Buchan KL, and Dunlop DJ (1975) Magnetization changes caused by burial and uplift. *Earth and Planetary Science Letters* 28: 133–143.

Purucker M and Whaler K (2007) Crustal magnetism. In: Schubert G (ed.) *Treatise on Geophysics: Geomagnetism*, vol. 5, pp. 195–237. Amsterdam: Elsevier.

Rayleigh L (1887) Notes on electricity and magnetism – III. On the behaviour of iron and steel under the operation of feeble magnetic forces. *Philosophical Magazine* 23(142): 225–245.

Reynolds RL, Rosenbaum JG, Hudson MR, and Fishman NS (1990) Rock magnetism, the distribution of magnetic minerals in the earth's crust and aeromagnetic anomalies. In: Hanna WF (ed.) *Geologic Applications of Modern Aeromagnetic Surveys: United States Geological Survey Bulletin 1924*, 24–45.

Roberts AP, Chang L, Rowan CJ, Horng C-S, and Florindo F (2011) Magnetic properties of sedimentary greigite (Fe_3S_4): An update. *Reviews of Geophysics* 49(1): RG1002, http://dx.doi.org/10.1029/2010rg000336.

Robinson P, Harrison RJ, Fabian K, and McEnroe SA (2012) Chemical and magnetic properties of rapidly cooled metastable ferri-ilmenite solid solutions: Implications for magnetic self-reversal and exchange bias – III. Magnetic interactions in samples produced by Fe–Ti ordering. *Geophysical Journal International* 191(3): 1025–1047. http://dx.doi.org/10.1111/j.1365-246X.2012.05692.x.

Robinson P, Harrison RJ, McEnroe SA, and Hargraves RB (2002) Lamellar magnetism in the haematite–ilmenite series as an explanation for strong remanent magnetization. *Nature* 418(6897): 517–520.

Robinson P, Harrison RJ, McEnroe S, and Hargraves RB (2004) Nature and origin of lamellar magnetism in the hematite–ilmenite series. *American Mineralogist* 89(5–6): 725–747.

Rochette P, Fillion G, Mattéi J-L, and Dekkers MJ (1990) Magnetic transition at 30–34 Kelvin in pyrrhotite: Insight into a widespread occurrence of this mineral in rocks. *Earth and Planetary Science Letters* 98: 319–328.

Rochette P, Jackson MJ, and Aubourg C (1992) Rock magnetism and the interpretation of anisotropy of magnetic susceptibility. *Reviews of Geophysics* 30: 209–226.

Rochette P, Mathé P-E, Esteban L, et al. (2005) Non-saturation of the defect moment of goethite and fine-grained hematite up to 57 Teslas. *Geophysical Research Letters* 32. http://dx.doi.org/10.1029/2005GL024196.

Rochette P, Weiss BP, and Gattacceca J (2009) Magnetism of extraterrestrial materials. *Elements* 5: 223–228.

Sagnotti L, Macrì P, Egli R, and Mondino M (2006) Magnetic properties of atmospheric particulate matter from automatic air sampler stations in Latium (Italy): Toward a definition of magnetic fingerprints for natural and anthropogenic PM10 sources. *Journal of Geophysical Research* 111(B12S22). http://dx.doi.org/10.1029/2006JB004508.

Schlinger CM, Veblen DR, and Rosenbaum JG (1991) Magnetism and magnetic mineralogy of ash flow tuffs from Yucca Mountain, Nevada. *Journal of Geophysical Research B: Solid Earth* 96: 6035–6052.

Schmidt V, Gunther D, and Hirt AM (2006) Magnetic anisotropy of calcite at room-temperature. *Tectonophysics* 418(1–2): 63–73.

Schmidt A, Richard Y, Matt H, and Ashmore M (2005) Magnetic susceptibility as proxy for heavy metal pollution: A site study. *Journal of Geochemical Exploration* 85: 109–117.

Shaar R, Ron H, Tauxe L, et al. (2010) Testing the accuracy of absolute intensity estimates of the ancient geomagnetic field using copper slag material. *Earth and Planetary Science Letters* 290(1–2): 201–213. http://dx.doi.org/10.1016/j.epsl.2009.12.022.

Shcherbakov VP and Fabian K (2005) On the determination of magnetic grain-size distributions of superparamagnetic particle ensembles using the frequency

dependence of susceptibility at different temperatures. *Geophysical Journal International* 162(3): 736–746.
Shcherbakov V and Sycheva N (2010) On the mechanism of formation of depositional remanent magnetization. *Geochemistry, Geophysics, Geosystems* 11(2). http://dx.doi.org/10.1029/2009gc002830.
Shimizu H, Koyama T, Koyama S, and Utada H (2007) A geomagnetic total intensity anomaly originated from lightning-induced isothermal remanent magnetization: Case of the Yatsugatake Magnetic Observatory, central Japan. *Earth, Planets and Space* 59(3): 141–149.
Spassov S, Egli R, Heller F, Nourgaliev DK, and Hannam J (2004) Magnetic quantification of urban pollution sources in atmospheric particulate matter. *Geophysical Journal International* 159(2): 555–564.
Strzyszcz Z and Magiera T (1998) Magnetic susceptibility and heavy metals contamination in soil of Southern Poland. *Physics and Chemistry of the Earth* 23(9–10): 1127–1131.
Syono Y (1960) Magnetic susceptibility of some rock forming silicate minerals such as amphiboles, biotites, cordierites and garnets. *Journal of Geomagnetism and Geoelectricity* 11(3): 85–93.
Tan XD, Kodama KP, Gilder S, and Courtillot V (2007) Rock magnetic evidence for inclination shallowing in the Passaic Formation red beds from the Newark basin and a systematic bias of the Late Triassic apparent polar wander path for North America. *Earth and Planetary Science Letters* 254(3–4): 345–357.
Tarduno JA and Wilkison SL (1996) Non-steady state magnetic mineral reduction, chemical lock-in, and delayed remanence acquisition in pelagic sediments. *Earth and Planetary Science Letters* 144: 315–326.
Tauxe L (2005) Inclination flattening and the geocentric axial dipole hypothesis. *Earth and Planetary Science Letters* 233: 3–4.
Tauxe L (2010) *Essentials of Paleomagnetism*. Berkeley University of California Press.
Tauxe L, Kodama KP, and Kent DV (2008) Testing corrections for paleomagnetic inclination error in sedimentary rocks: A comparative approach. *Physics of the Earth and Planetary Interiors* 169: 152–165.
Tauxe L, Steindorf JL, and Harris A (2006) Depositional remanent magnetization: Toward an improved theoretical and experimental foundation. *Earth and Planetary Science Letters* 244(3–4): 515–529.
Tauxe L and Yamazaki T (2007) Paleointensities. In: Schubert G (ed.) *Treatise on Geophysics, Geomagnetism*, pp. 509–564. Oxford: Elsevier Ltd.
Thompson R and Oldfield F (1986) *Environmental Magnetism*. London: Allen and Unwin.
Till JL, Jackson MJ, Rosenbaum JG, and Solheid P (2010) Magnetic properties in an ashflow tuff with continuous grain size variation: A natural reference for magnetic particle granulometry. *Geochemistry, Geophysics, Geosystems* 12(Q07Z26). http://dx.doi.org/10.1029/2011GC003648.
Tite MS and Mullins CE (1971) Enhancement of the magnetic susceptibility of soils on archaeological sites. *Archaeometry* 13(2): 209–219.
Turner GM, Rasson JL, and Reeves CV (2007) Observation and measurement techniques. In: Schubert G (ed.) *Treatise on Geophysics*, vol. 5, pp. 93–146. Amsterdam: Elsevier.
Verosub KL and Roberts AP (1995) Environmental magnetism: Past, present, and future. *Journal of Geophysical Research B: Solid Earth* 100: 2175–2192.
Verrier V and Rochette P (2002) Estimating peak currents at ground lightning impacts using remanent magnetization. *Geophysical Research Letters* 29(18): 14-1–14-4. http://dx.doi.org/10.1029/2002GL015207.
Walz F (2002) The Verwey transition – A topical review. *Journal of Physics: Condensed Matter* 14(12): R285–R340.
Williams W and Dunlop DJ (1989) Three-dimensional micromagnetic modelling of ferromagnetic domain structure. *Nature* 337: 634–637.
Williams W and Muxworthy AR (2006) Understanding viscous magnetization of multidomain magnetite. *Journal of Geophysical Research: Solid Earth* 111(B2).
Wolfers P, Fillion G, Ouladdiaf B, Ballou R, and Rochette P (2011) The Pyrrhotite 32 K magnetic transition. *Solid State Phenomena* 170: 174–179.
Worm H-U (1991) Multidomain susceptibility and anomalously strong low field dependence of induced magnetization in pyrrhotite. *Physics of the Earth and Planetary Interiors* 69: 112–118.
Worm H-U, Clark D, and Dekkers MJ (1993) Magnetic susceptibility of pyrrhotite: Grain size, field and frequency dependence. *Geophysical Journal International* 114: 127–137.
Worm H-U and Jackson M (1999) The superparamagnetism of Yucca Mountain Tuff. *Journal of Geophysical Research B: Solid Earth* 104(B11): 25415–25425.
Yamazaki T (2008) Magnetostatic interactions in deep-sea sediments inferred from first-order reversal curve diagrams: Implications for relative paleointensity normalization. *Geochemistry, Geophysics, Geosystems* 9(Q02005). http://dx.doi.org/10.1029/2007GC001797.
Yaramanci U and Müller-Petke M (2013) Tools and techniques: Nuclear magnetic resonance. In: Slater L (ed.) *Treatise on Geophysics*, vol. 11, pp. 93–146. Elsevier.
Yu YJ and Tauxe L (2006) Acquisition of viscous remanent magnetization. *Physics of the Earth and Planetary Interiors* 159(1–2): 32–42.

11.06 Tools and Techniques: Marine Seismic Methods

JOA Robertsson, ETH-Zurich, Zürich, Switzerland
R Laws and E Kragh, Schlumberger Gould Research, Cambridge, UK

© 2015 Elsevier B.V. All rights reserved.

11.06.1 Introduction 175
11.06.2 The Seismic Method 177
11.06.2.1 Reflection and Refraction Seismic Methods 177
11.06.2.2 Velocity Models and Imaging 178
11.06.2.3 Survey Objectives and Designs 178
11.06.3 Data and Noise 179
11.06.3.1 Residual Shot Noise 180
11.06.3.2 Seismic Interference 180
11.06.3.3 Ambient Noise, Tow Noise, and Group Forming 180
11.06.3.4 Perturbations and 'Noise Behind the Signal' 181
11.06.3.5 Incorrectly Imaged Signal 181
11.06.3.6 Ghosts and Multiples 181
11.06.3.7 Seabed Seismic-Specific Issues 184
11.06.3.7.1 Scholte waves 184
11.06.3.7.2 Vector fidelity 184
11.06.4 Acquisition Systems 184
11.06.4.1 Marine Recording Systems 185
11.06.4.2 Seabed Recording Systems 186
11.06.4.3 Marine Sources: Air Guns 186
11.06.4.4 Marine Sources: Vibrators 192
11.06.4.5 Steering, Navigation, and Positioning 193
11.06.5 Illumination 194
11.06.5.1 Seismic Survey Configurations Geometries 194
11.06.5.2 Seismic Sources 194
11.06.5.2.1 Single-source and dual-source flip–flop 194
11.06.5.2.2 Simultaneous sources 194
11.06.5.2.3 Azimuthally invariant sources 195
11.06.6 Sampling and Spatial Aliasing 195
11.06.6.1 Receiver-Side Techniques 197
11.06.6.2 Source-Side Techniques 198
11.06.7 Bandwidth 198
11.06.7.1 Receiver-Side Techniques 199
11.06.7.1.1 Single-streamer deghosting methods 199
11.06.7.1.2 Over–under methods 200
11.06.7.1.3 Variable tow depth methods 200
11.06.7.1.4 Multicomponent towed streamer methods 200
11.06.7.1.5 Seabed seismic techniques 201
11.06.7.2 Source-Side Techniques 201
11.06.8 Time-Lapse and Reservoir Monitoring 201
11.06.9 Conclusions 203
Acknowledgments 205
References 205

11.06.1 Introduction

The seismic method was originally pioneered in land environments. However, as hydrocarbon exploration moved offshore, so did the seismic method. Having started off as a two-dimensional (2D) method with 10s of recording channels from sensors in a single 'streamer,' most marine seismic data are today acquired using 3D acquisition systems. Figure 1, where a marine seismic spread has been superimposed on top of the city of Oslo in the upper panel, serves to illustrate the main components and the size of a conventional 3D marine seismic acquisition system. Such acquisition systems have been referred to as the largest moving man-made objects on Earth: The dimensions typically measure 1–1.5 km across and 6–8 km in length. At the front of the spread, the seismic vessel is located. The vessel tows sources and receivers (located

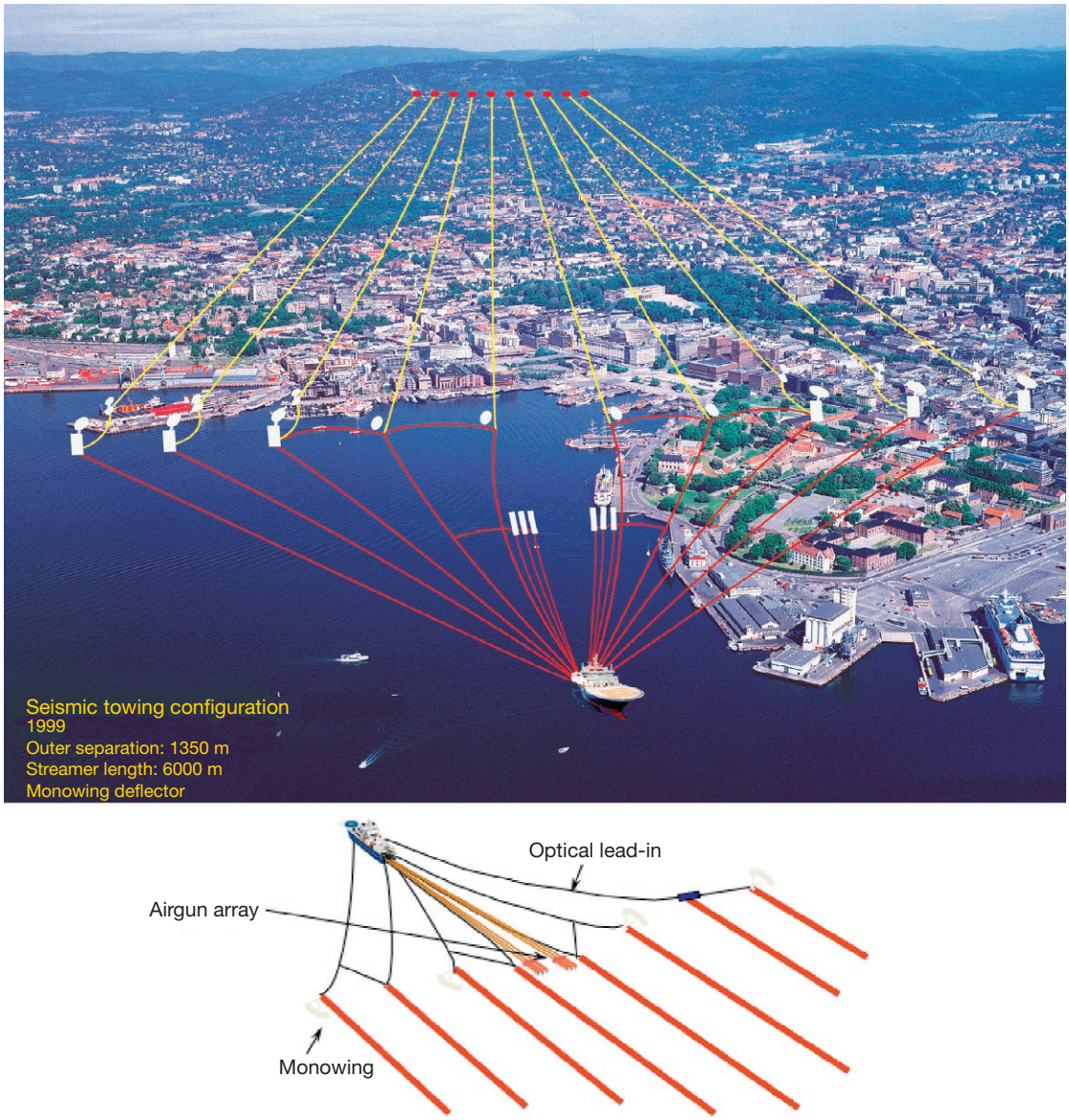


Figure 1 Top: Conventional 3D seismic spread overlain on the city of Oslo, Norway, for illustration of scale (photo provided courtesy of WesternGeco). Bottom: Some detail of the front-end architecture.

in streamers) through the water. Data are also recorded and sources and streamers controlled and steered from the vessel and beneath the water surface. Straight behind the vessel, one or two marine seismic sources are usually towed emitting seismic energy into the subsurface. The reflections from the Earth's subsurface are recorded using sensors located along streamers towed behind the vessel. In the top panel of Figure 1, the acquisition system comprises 10 streamers spaced 150 m apart. Marine seismic methods concern not only towed marine seismics but also acquisition systems using sensors deployed on the seabed using cables or nodes. We will also cover such acquisition methods in this article.

Figure 2 shows a typical state-of-the-art example of an image obtained from 3D marine seismic data acquired for hydrocarbon exploration. The quality of marine data tends to be very good compared to land seismic equivalent examples. From images such as the one in Figure 2, the interpretation of subsurface structures of geologic significance can be made (e.g., unconformities, faults, salt bodies, and meandering paleoriver channels). Such information can reveal whether hydrocarbon systems have developed and whether there are likely traps for hydrocarbon exploration. In this chapter, we will be reviewing the acquisition systems and methods leading to marine seismic images such as these.

The marine seismic method is not only used for hydrocarbon exploration applications. Other applications include shallow imaging for mineral exploration and engineering site investigation surveys (e.g., Müller et al., 2009), deep crustal imaging (e.g., Singh et al., 2008), or seismic oceanography (Holbrook et al., 2009). However, in this chapter, we will be

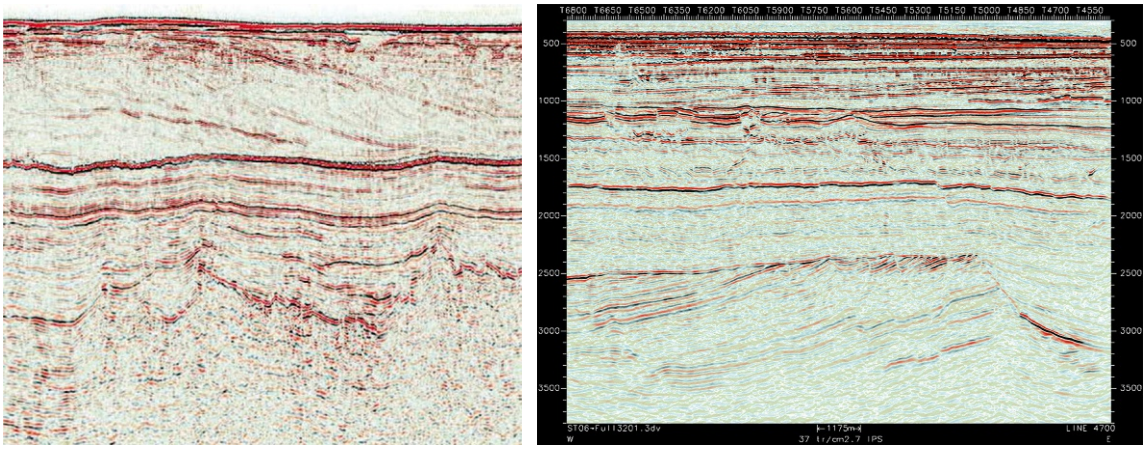


Figure 2 Examples of vertical sections from final seismic images following processing and migration of the acquired 3D marine seismic data. The image to the left is from the Buzzard field in the North Sea (provided courtesy of WesternGeco), whereas the image to the right is from the Snorre field also in the North Sea (provided courtesy of Statoil). Images such as these enable a detailed geological interpretation of sub-surface structures. The dimensions are a few km both laterally and vertically.

focusing on describing the method from a hydrocarbon exploration perspective since it is the most advanced application area and often drives the evolution of other marine seismic applications.

We will start with a review of the seismic method and possible objectives behind acquiring and imaging seismic data. Although this chapter is not about data processing, we will need to briefly discuss how data are processed to appreciate the impact that it has on how data are acquired. Next, we will discuss various types and sources of noise that impact towed marine and seabed seismic data. We then describe the acquisition systems: towed marine streamers, seabed systems, marine seismic sources, and technology used for steering, navigating, and positioning the acquisition systems. The next three sections will discuss three fundamental issues governing the quality of seismic data and their ability to create an image of the subsurface: illumination of the subsurface, spatial sampling, and frequency bandwidth. Finally, in most of the chapter, we will be discussing the seismic method from an exploration seismic perspective. Therefore, at the end, we will discuss time-lapse and reservoir monitoring applications where repeat seismic surveys are conducted, for instance, to monitor the migration of fluid fronts in a reservoir during production. Such applications put different constraints on how the data are acquired and processed compared to more conventional exploration applications.

11.06.2 The Seismic Method

In a seismic survey, a source is excited, generating waves that propagate into the subsurface. As these waves encounter subsurface heterogeneities, they will undergo reflections, transmissions, refractions, and conversions into other wave types (P-waves, S-waves, interface waves, etc.). Some of this energy will reverberate back to the surface where seismic signals are recorded at a large number of receiver locations. The resulting data are referred to as a shot record. Once all energy from the shot has been recorded, a new shot record is collected from a different source location. In this fashion, a seismic data set is gradually built up by sequentially exciting shots at different locations.

There are two types of seismic methods: the reflection seismic method and the refraction seismic method. The discussion in the chapter related to processing techniques, noise considerations, and survey design will mostly be based around the reflection seismic method since it completely dominates resource exploration applications.

11.06.2.1 Reflection and Refraction Seismic Methods

In the refraction seismic method, we will focus on the kinematics of waves that are either continuously refracted as diving waves or critically refracted at interfaces and returned to the surface as head waves to determine the velocity structure and the depth and dips of the major discontinuities. The refraction seismic method will naturally require long offsets as most diving wave refractions and head waves in particular are only generated beyond the angle of critical reflection at major subsurface interfaces. Critical refraction only occurs for those interfaces across which the seismic velocity increases.

In the reflection seismic method on the other hand, the focus is on the shorter offsets (or angles of incidence) up until critical reflections. The objective of seismic reflection data processing is to produce an image of the subsurface. The acquired seismic data are processed to isolate the primary reflections from subsurface interfaces, whereas all other arrivals (surface waves, head waves, multiple reflections, etc.) are considered as noise that needs to be removed before imaging the seismic data.

A central concept in the reflection seismic method is the common midpoint (CMP), which simply corresponds to the horizontal position halfway between the source location and the receiver location for a recorded data trace. Under the assumption that the Earth is horizontally layered and that the elastic properties of each layer are either isotropic or anisotropic with a vertical symmetry axis, the CMP corresponds to the horizontal location at which all primary reflections on the recorded trace are reflected (at least under a high-frequency or

ray-based assumption where reflections take place at a point). Therefore, traditional reflection seismic data processing workflows have been based around processing data sorted into CMP gathers loosely following a workflow as follows:

1. Determine the velocity profile (as a function of recorded two-way travel time) at the CMP location from the curvature of the reflections (moveout) in the CMP gather and their travel time at zero offset.
2. Apply a normal-moveout (NMO) correction to the CMP gather using the velocity profile to align the travel time of traces to that of the zero-offset trace.
3. Stack (sum) all traces in the CMP gather.
4. Migrate (image) the stacked data at all CMPs to produce an image where subsurface reflectors have the correct horizontal position, dip, and two-way travel time (corresponding to normal incidence that is also assumed to be the vertical propagation path).

The workflow mentioned earlier is an example of a poststack time migration method and is rarely used in practice today since it is not sufficiently accurate for lateral variations in the subsurface. Nevertheless, the concept of the CMP is still remarkably robust and still has a central role also in many prestack imaging workflows; today, most migration methods are applied in this prestack domain. In this context, the importance of ensuring a uniform distribution of offsets on different CMP gathers is one of the most critical parameters when choosing source and receiver locations during the design of a seismic survey. A uniform distribution of offsets between CMPs results in better structural continuity in the image and the ability to estimate elastic properties of the subsurface.

11.06.2.2 Velocity Models and Imaging

A fundamental assumption underlying the classic reflection seismic method is that the macro(long-wavelength)-velocity model can be separated from the reflectivity distribution in the subsurface. The reflections originate from the short-wavelength variations in the impedance (product of density and velocity) that controls the amplitudes of the reflections, but the travel times are mainly determined by the long-wavelength velocity model. A number of different imaging methods exist that all assume that a velocity model has first been determined either as a function of two-way travel time or as a function of depth. A good summary of recent developments in velocity estimation for depth imaging (reflection tomography, diving wave tomography, waveform inversion, etc.) can be found among the papers presented in a recent special section on the subject in the journal *Geophysics* (Harlan et al., 2008; Woodward et al., 2008).

The literature on imaging methods is huge. In the 1990s, Kirchhoff prestack depth migration was the workhorse for high-end imaging in the oil industry (Bleistein et al., 2001). The method is based on a high-frequency ray assumption that may be unsuitable in complex media with sharp velocity contrasts and rough interfaces and structures. Instead, Gaussian beam methods were developed to address some of these shortcomings (Hill, 2001). The real breakthrough in imaging in complex media came in the early to mid-2000s with the successful application of full waveform imaging methods such as one-way wave-equation migration (OWEM) and reverse time migration (RTM). Etgen et al. (2009) presented an excellent review of the evolution and state-of-the-art of depth imaging methods.

The breakthrough in imaging in complex media came both from the mentioned advances in imaging methods and from the introduction of marine acquisition techniques that better illuminated the subsurface target. The evolution was driven by the need for better imaging under complex salt structures in the Gulf of Mexico. Through advanced illumination studies, it became clear that traditional narrow-azimuth 3D marine seismic data did not capture sufficient ray paths illuminating the target (Regone, 2007; VerWest and Lin, 2007). Instead, the so-called wide-azimuth (WAZ) acquisition techniques were introduced (these are discussed further in **Section 11.06.5.1**). However, WAZ acquisition geometries did not produce data that were easily sorted into gathers suitable for Kirchhoff migration. Instead, the OWEM and RTM methods naturally operate on source or receiver gathers (data traces with a common shot or receiver location) that more naturally fit the data acquired in WAZ surveys.

In summary, fundamental changes both in the way that marine seismic data are acquired and in the methods used for imaging have occurred over the last decade. The resulting images are significantly improved and are capable of, for example, resolving many complex subsalt plays for the first time (e.g., Vigh et al., 2011). **Figure 3** shows a direct comparison of the same image produced using state-of-the-art acquisition and imaging technology in 2005 versus 2010. The improvement in the subsalt image follows from the combination of new acquisition and imaging techniques.

Finally, the split between velocity and reflectivity assumed in classic seismic reflection data processing can be avoided by taking a different approach to subsurface imaging and inversion; the full waveform inversion (FWI) method was introduced in the 1980s by Tarantola and coworkers (e.g., Tarantola, 1984). FWI is an iterative method where seismic data are synthesized (modeled) using, for instance, a finite-difference method on an initial earth model. By designing a cost function that attempts to minimize differences between modeled and real data, the FWI method provides an update to the Earth model that then can be used to generate new synthetic data. After several iterations, the FWI method provides an Earth model that is increasingly capable of generating a better fit between synthetic and real data. The FWI method is highly computationally expensive and did therefore not find much practical use until the 2000s (Pratt and Brenders, 2004; Virieux and Operto, 2009). It is also highly dependent on a good initial model in order to converge and avoid becoming trapped in local minima. However, recently, the method has been successfully applied in wide range of hydrocarbon exploration projects (e.g., Sirgue et al., 2011; Vigh et al., 2011).

11.06.2.3 Survey Objectives and Designs

Depending on the objective or geologic setting, the design of the seismic survey will vary. If the primary purpose of the survey is an imaging objective for exploration, then illumination of the subsurface target and perhaps addressing known

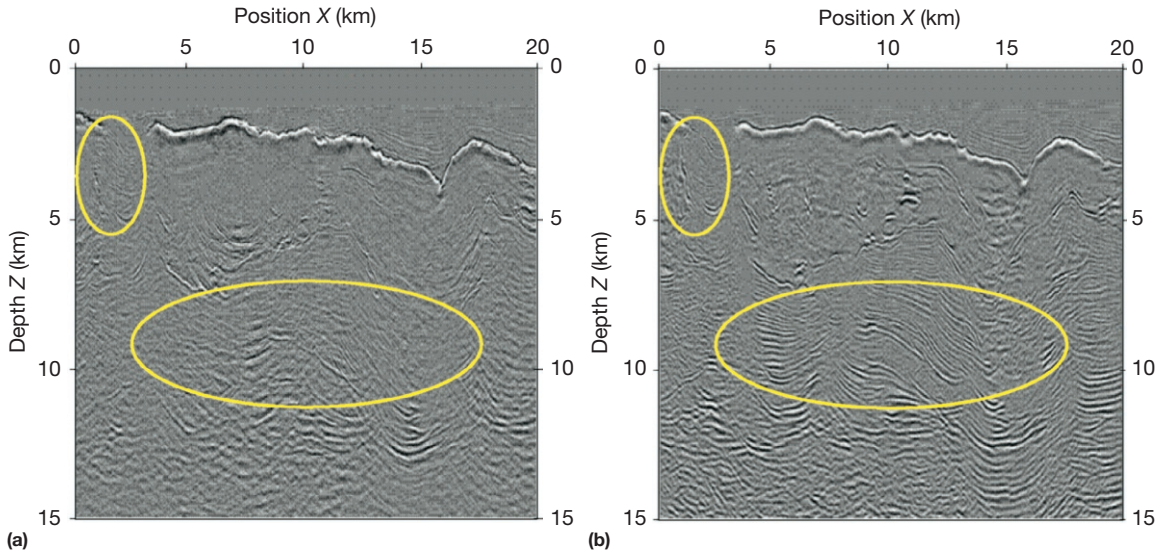


Figure 3 (a) Prestack depth migrated image using narrow-azimuth data and OWEM from 2005. (b) Prestack depth migrated image using wide-azimuth data and RTM from 2010. Reproduced from Vigh D, Kapoor J, Moldoveanu N, and Li H (2011) Breakthrough acquisition and technologies for subsalt imaging. *Geophysics* 76: WB41–WB51.

geophysical noise problems in the area will take precedence. If the survey has a time-lapse seismic objective for reservoir monitoring (see Section 11.06.8), we need to make sure that the data are highly repeatable between consecutive surveys and that noise levels are kept to a minimum.

When designing a seismic survey meeting the desired objectives, an approximate subsurface model in terms of velocities (and possibly major reflecting interfaces such as salt structures) is needed. In order to make sure that sufficiently densely acquired data with uniform illumination of the target are acquired, the anticipated data traces are typically assigned to 'bins' corresponding to their CMP location. All traces belonging to the same bin will be considered as having the same CMP. The bin dimensions are chosen to obtain the required and achievable lateral resolution of the target so that the data can be migrated properly. Once all desired source/receiver pairs for the entire survey have been mapped out and sorted into bins, one should make sure that the bins are uniformly sampled in terms of both offsets and the number of data traces. This is to ensure that the data can be processed without leaving imprints of the survey geometry in the imaged data (the so-called acquisition footprint).

In designing the seismic survey, one also needs to bear in mind the desired data processing workflow. Several of the data processing steps that may be required will have specific and sometimes different survey design criteria. For example, if surface-related multiples are known to be a problem (see Section 11.06.3.6), then a uniform sampling of short offsets will be required in order for surface-related multiple elimination (SRME) methods to work (Dragoset and Jericevic, 1998). If we expect interfering energy from other seismic surveys (known as seismic interference; see Section 11.06.3.2) to be a problem, then we might need to choose the sail lines to be in line with the dominant directions of the seismic interference so that it can be easily removed. Finally, if we are dealing with a complex subsalt play, we may need to use RTM (see Section 11.06.2.2) to image the target. Whether the data can be easily sorted into common-source or common-receiver gathers will be a particularly important question in such a scenario.

In complex imaging, sorting data into CMP bins to assess the survey is typically not sufficient. Instead, it will be necessary to numerically model the data using, for instance, finite-difference methods and an anticipated subsurface model (Regone, 2007) to make sure that a survey will acquire all data that will produce a high-quality image. This sounds like a chicken-and-egg problem since we do not exactly know what the subsurface model looks like before acquiring and imaging the data. However, usually, a rough idea of the anticipated subsurface structure is available that should be sufficient to design an appropriate survey strategy.

Finally, when planning a seismic survey, it is also worthwhile to 'future proof' the survey beyond the immediate survey objectives as seismic data tend to be reused for different purposes in the future. Particular questions to bear in mind include the following: Will the data be easy to repeat in a future time-lapse survey? Do the data contain short or long offsets required for multiple attenuation or seismic inversion?

11.06.3 Data and Noise

Historically, it was simple to think of primary reflections (i.e., that energy that propagates down from the source and is reflected directly back to the receivers) as our data, or signal, and everything else as noise. Notable exceptions to this are in land data with the use of shallow refraction surveys to build near-surface velocity models and long-offset refraction studies to build deeper velocity models (the refraction seismic method was briefly described in Section 11.06.2.1). Over the years, as the seismic method has progressed, more and more of the shot-generated wavefield is being considered 'signal'

(i.e., energy that we can derive some useful information from). For example, though again on land surveying, the use of surface wave energy (traditionally the highest amplitude 'noise' in land surveys) is now routinely used to derive near-surface model information (Strobbia et al., 2011). Multiple reflection energy (see Section 11.06.3.6) was always considered 'noise' in that we try to remove it from our data. However, methods now exist to use these multiples to improve the image (e.g., Berkhout and Verschuur, 2012; Muijs et al., 2007a,b). Ultimately, elastic FWI should be able to use all the shot-generated seismic wavefield to obtain an image (Virieux and Operto, 2009). However, having mentioned all this, even with current state-of-the-art seismic processing, the bulk of work looks to image the primary reflection signal, and we are back to our opening statement: this is our data, and everything else is noise. The following sections consider some of the typical noise types that are observed in marine seismic data.

11.06.3.1 Residual Shot Noise

In marine seismic acquisition, the time between consecutive shots being fired is typically around 10 s. With a typical vessel speed of 2.5 m s^{-1} (about five knots), this corresponds to a distance of 25 m. There are a number of factors that contribute to this time between shots. Firstly, we want to record reflection data from the target reflectors that arrive after a few seconds. In addition to this minimum recording time, we need to wait a little longer to record enough data in order for our imaging and migration algorithms to work correctly (the so-called migration time aperture). A second factor is that the compressors that charge the high-pressure air gun sources need time to charge. A third reason is that we need to wait for the energy from the shot to die down in amplitude to a small enough level, such that it does not interfere with energy from the next shot. This remaining energy from the previous shot is residual shot noise, and it is one component of the shot-generated noise. If we decrease the time between shots, the residual shot energy increases and may become problematic as it is challenging to separate this energy out. Figure 4 shows a typical shot record where the recording time is 12 s. The reflection signal may not be visible at the end of the record, but it still has a small remaining amplitude contributing to the noise field. The energy at the very start of the record contains the residual shot noise from the previous shot.

11.06.3.2 Seismic Interference

It is quite common for more than one seismic company to be acquiring surveys at the same time. While these surveys might be separated by many 10s or 100s of kilometers, the direct wave that travels through the water decays slowly in amplitude and can interfere with the neighboring survey. This noise, called seismic interference, can be sufficiently strong as to require the surveys to take turns on acquisition; this is referred to as time-sharing. Time-sharing can have a significant impact on the economics of the survey, but although there are methods to remove this interfering energy (e.g., Gulunay, 2008), it is typically difficult to attenuate and time-sharing remains standard practice.

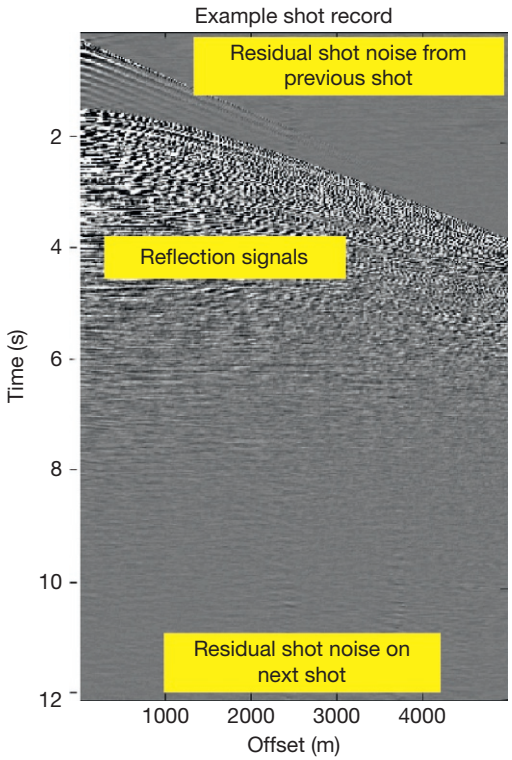


Figure 4 A typical marine seismic shot record with 12 s recording time shown. The signal strength decays with time, but when the next shot goes off, any residual energy remaining is called residual shot noise.

11.06.3.3 Ambient Noise, Tow Noise, and Group Forming

The ocean ambient noise is the background noise that is present irrespective of the seismic survey being present or not. The ambient noise is caused by waves, shipping, marine animal life, and general man-made activity. It is typically the smaller component of the total noise field in marine seismic surveying, although it can become significant at low frequencies below 10 Hz or so, where it rises in amplitude and coincides with the part of the trace where the seismic signal falls sharply in amplitude (Laws et al., 2008).

Tow noise is the dominant form of the noise caused by the acquisition system, but not generated from the seismic shot. It forms part of the non-shot-generated noise. Depending upon the physical construction of the streamer (solid, gel, or containing liquid), the tow noise may be physical strumming of the streamer under tension, 'swell noise' that is energy imparted to the streamer from the sea-surface waves, or 'flow noise' that is caused by turbulent fluctuations in the vicinity of the streamer. Elboth et al. (2009) give a good technical review. Figure 5 shows a typical shot record where the swell noise is clear at certain offsets (the high-amplitude bursts of vertical energy). Swell noise is reduced by towing the streamers deeper, which means they are further away from the effects of the sea surface, something which has been enabled by recent multicomponent streamer designs (see Section 11.06.7.1.4) and previously was enabled by 'over–under' acquisition methods (see Section 11.06.7.1.2). Again, Figure 5 shows this effect clearly. The streamer in the right-hand panel is towed at 20 m

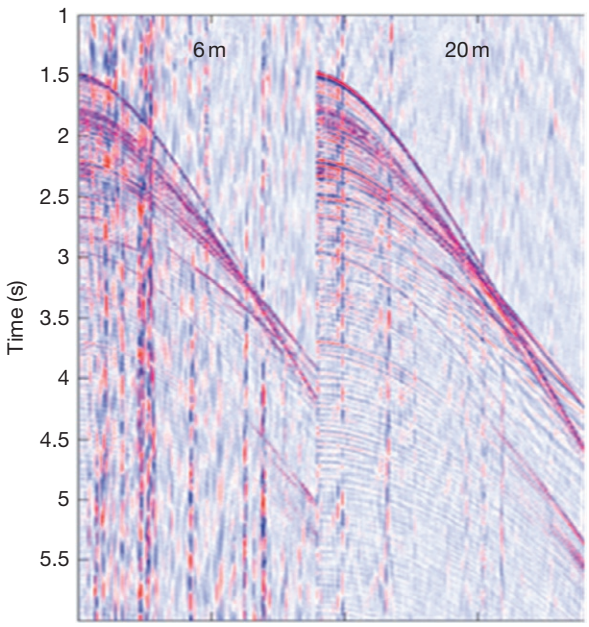


Figure 5 Example of shallow- and deep-tow shot records (2 Hz low-cut filter applied). The data acquired at 20 m tow depth clearly have both less noise and stronger (low-frequency) reflection signal. Reproduced from Kragh E, Muyzert E, Curtis T, Svendsen M, and Kapadia D (2010) Efficient broadband marine acquisition and processing for improved resolution and deep imaging. *The Leading Edge* 29: 464–469.

depth, while the streamer in the left-hand panel is towed at 6 m depth and shows much more swell noise contamination. Signal processing methods for swell noise removal are prevalent and form the basic first step in marine seismic processing and acquisition (e.g., Moldoveanu, 2012). However, the most basic method to attack swell (and ambient) noise is group forming.

Group forming is simply the summation of a number of closely spaced recording elements to form a single-output channel, and most marine seismic streamers are analog group formed, that is, the elements are hardwired together in the streamer. The length of the group is designed to be small compared with the high apparent wavelength of the near-vertically returning reflection signals but is typically longer than the apparent wavelength of the swell noise. As such, reflection signal components constructively interfere, but the noise components are averaged out; thus, the overall signal-to-noise ratio is increased on output. Typical group lengths are between 12 and 25 m in modern analog streamers, and this defines the output spacing of the seismic measurement from the streamer. Even though the group lengths are short compared to the signal apparent wavelengths, high-frequency signal and high-dipping signal can be attenuated using analog groups particularly if perturbations (errors) are present with the group (see Section 11.06.3.4). Finally, there is one class of streamers that record data separately from all the closely spaced recording elements without performing analog grouping. These point-receiver recording streamers allow for digital group forming to be performed after the data are recorded, and this affords more flexible and powerful noise attenuation (e.g., see Martin et al., 2000). The design and architecture of marine seismic streamers is discussed further in Section 11.06.4.1.

11.06.3.4 Perturbations and 'Noise Behind the Signal'

Perturbations are essentially the errors on the individual measurements. Examples of such perturbations include uncalibrated sensors and rough sea perturbations that vary from sensor to sensor. When the individual measurements are summed together without correcting for these perturbations, for example, during analog group forming or later in the processing using multichannel filters, an underlying distortion is imparted on the data that can be difficult to remove. This is termed 'noise behind the signal.' It is the total error in the signal fidelity and limits the dynamic range of the data and the ability to distinguish and image weak-amplitude events. This type of error can be particularly important for time-lapse seismic (see Section 11.06.8) where small changes in the seismic signal over time are imaged. Laws and Kragh (2002) discussed in detail a particular perturbation and its impact on the time-lapse image. This is the effect of the rough sea that imparts a perturbation on the ghost reflection (the ghost reflection is described in Figure 6 and Sections 11.06.3.6 and 11.06.7.1) in the form of an amplitude and timing error. Uncorrected, this small error can give rise to a false time-lapse signal, confounding a proper interpretation.

11.06.3.5 Incorrectly Imaged Signal

By incorrectly imaged signal, we mean seismic data originating from a shot that in principle could contribute to an image of the subsurface but that has not been processed appropriately for reasons as follow:

- Residual shot noise (Section 11.06.3.1) could contribute to the image if this noise were to be associated with its corresponding shot.
- Nonreflected seismic arrivals such as refractions and Scholte waves (Section 11.06.3.7.1) sense subsurface properties but will not be properly handled in a reflection seismic processing workflow.
- Using an incorrect velocity model will incorrectly position reflectors.
- The reflection seismic workflow assumes primary reflections only. Multiple reflections (Section 11.06.3.6) are often strong in recorded seismic data and could in principle provide important subsurface information.

11.06.3.6 Ghosts and Multiples

Traditionally, seismic data processing methods assume that all data consist of primary reflections, that is, energy emitted from the source reflects at one single interface in the subsurface and is recorded at a receiver when reaching the surface of the Earth again. In reality, waves also follow many other different propagation paths. In particular, sea-surface ghosts have a profound impact on marine seismic data (Figure 6). The source ghost corresponds to energy that is radiated upward from the source and immediately reflects downward to interfere with the

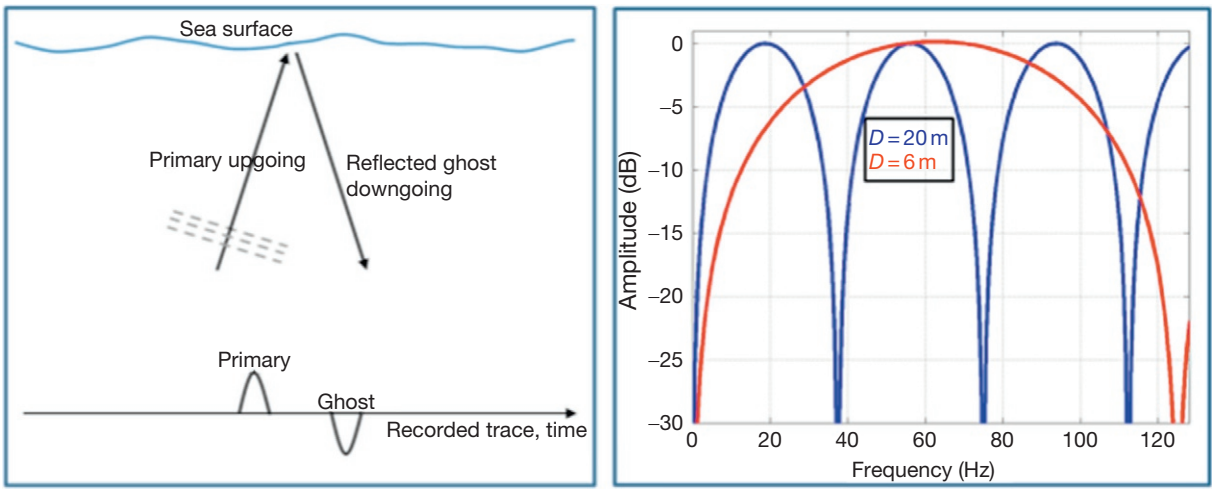


Figure 6 Sea-surface 'ghost' reflection concept and examples of resulting spectra for vertical incidence at 6 and 20 m depth. A ghost is generated both at the source and receiver compounding the effect on bandwidth (see, e.g., **Figure 5**).

emitted wavefield that is directly radiated downward. Similarly, the receiver ghost corresponds to subsurface reflections that reflect from the sea surface downward, interfering with the recorded upgoing wavefield. The main effect of the source and receiver ghosts is to severely reduce the bandwidth of the seismic data (Amundsen, 1993b), and we therefore discuss the ghost effect in more detail in **Section 11.06.7**. However, in addition, if the sea surface is rough, the ghosts will introduce static shifts on the source side and receiver side as well as noise due to scattering of the seismic wavefield from the sea surface (Laws and Kragh, 2002). Noise due to rough-sea-surface scattering can be of particular importance in time-lapse seismic applications (Laws and Kragh, 2002).

Multiples correspond to energy that reflects several times in opposite directions before being recorded at the surface and are classified according to how they reflect at interfaces before arriving at the receiver; this is illustrated in the upper panel of **Figure 7**. One of the most important categories of multiples in marine seismics is referred to as surface-related multiples. These are multiples that reflect from the sea surface at least once before being recorded (excluding the source and receiver ghosts). For surface-related multiples, the sea surface is referred to as the multiple-generating horizon. In the upper panel of **Figure 7**, we also illustrate three different types of surface-related multiples that can be of different orders (first order, second order, third order, etc.) depending on how many downward reflections from the sea surface they undergo. A real data example of a first-order surface-related multiple is shown in the lower panel of **Figure 7**.

An interbed multiple, on the other hand, is a multiple that has subsurface interfaces as multiple generating horizons instead of the sea surface (**Figure 7**). Multiples can of course also comprise both parts, namely, surface-related parts and those of interbed character.

Finally, a special case of a surface-related multiple that often is particularly prominent in marine seismic data is the so-called water-layer peg legs, which correspond to energy that reverberates several times within the water column on either the source side (source-side peg legs) or the receiver side (receiver-side peg legs).

The effect of multiples in data processing varies depending on the severity of the multiple problem and the subsurface complexity. Many times, the multiples can be clearly recognized as a time-delayed 'copy' of an earlier primary (as in the lower panel of **Figure 7**). However, other times, the multiple can be very difficult to recognize and leads to interpretation of false structure in the subsurface. The so-called scattered multiples are particularly difficult to deal with. These are multiples that arrive at the receiver (or emerge from the source) with a considerable cross sail-line propagation component. Since spatial sampling in the cross-line direction is often poor, multiple attenuation methods tend to be somewhat ineffective in handling these types of multiples. Finally, improperly attenuated multiples can also severely affect the signal-to-noise ratio after imaging.

Multiple removal methods can broadly be characterized as belonging to one of three different categories, namely, methods based on separability (moveout) and periodicity or methods based on modeling (wave-equation-based prediction and subtraction).

The most common multiple attenuation method based on separability often involves transformation of the data using the so-called parabolic Radon transform (Sheriff and Geldart, 1995). The transform maps data from the offset-time domain into a domain of parabolic curvature versus intercept time of events at zero offset. Since multiples tend to have a greater degree of curvature compared to primary reflections at similar arrival times (due to the fact that the multiples usually have propagated in shallower parts of the subsurface where propagation velocities are lower), they will often separate well from primary reflections after the parabolic Radon transform. The effect can also be enhanced by applying an NMO correction with primary propagation velocities before the Radon transform. After surgically isolating the multiples from the signal, the isolated multiples can be inversely transformed back to the offset-time domain and adaptively subtracted from the original data.

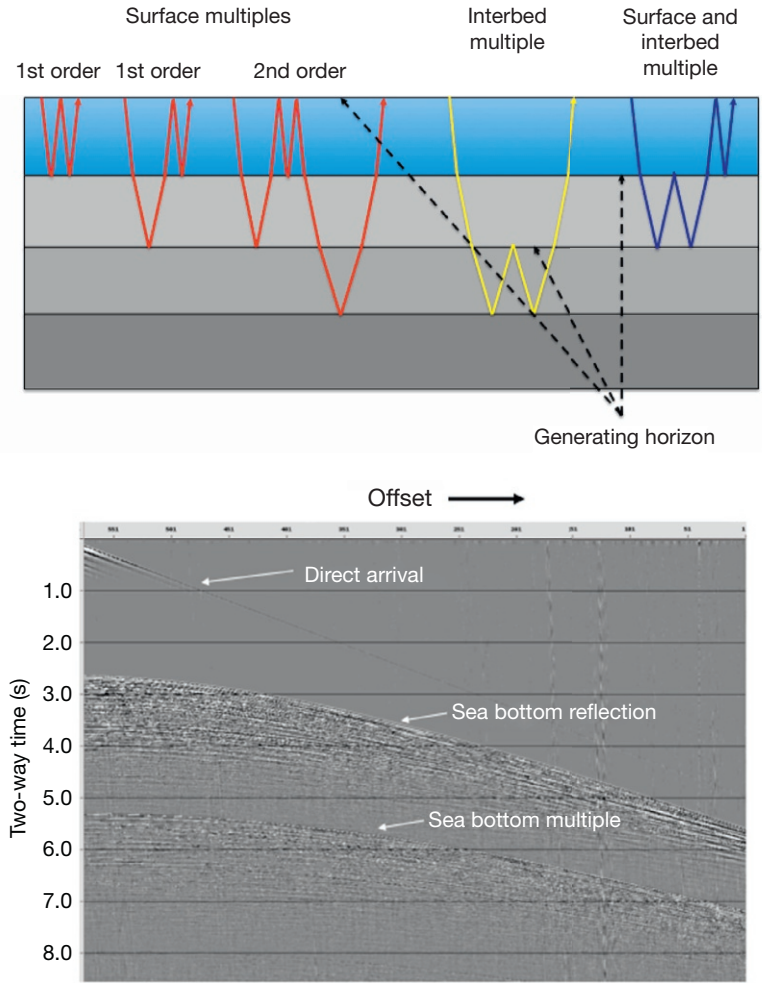


Figure 7 Upper panel: Classification of multiples in different categories. Lower panel: Example of appearance of a first-order surface-related multiple in a deep-water shot record.

Multiple attenuation methods based on periodicity involve the linear Radon transform where data in the offset-time domain are mapped into the slowness–zero-offset intercept time domain (often referred to as the tau-p domain; Sheriff and Geldart, 1995). After such a transformation, multiples can be removed through predictive deconvolution of data traces corresponding to different slowness (or p) values.

Finally, in multiple attenuation methods based on numerical forward modeling, a model of the multiples is first produced before (adaptively) subtracting it from the data. The most common and highly successful method for surface-related multiple attenuation is referred to as SRME (Surface-Related Multiple Elimination) and was originally developed by researchers at Technical University of Delft in the Netherlands (Dragoset and Jericevic, 1998; Verschuur et al., 1992). The method is based on the observation that any surface-related multiple can be regarded as being composed of separate legs corresponding to individual primary reflections. Through a spatial convolution of the recorded primaries with themselves (over both source and receiver coordinates), a model of surface-related multiples can be produced without any knowledge of the subsurface structure. The method can be highly effective but relies on adequate sampling of data on the surface, particularly for short source–receiver offsets. In its full 3D formulation, the method can become highly computationally intensive due to the large number of convolutions and interpolation between different source and receiver locations that is required.

Another multiple suppression method based on modeling is the wave-equation modeling method based on the so-called one-way wave equation where multiples are modeled using a synthetic model of background velocities and reflectors (Stork et al., 2006).

Finally, Amundsen (2001) introduced yet another multiple elimination method that relies on the decomposition of the recorded wavefield into its up- and downgoing constituents. The downgoing wavefield can be deconvolved from the upgoing wavefield removing both the surface-related multiples and the source signature. Decomposition of the recorded data is closely analogous to deghosting the data (see **Section 11.06.7**) and will typically require multicomponent recordings. Until recently, this was only available for seabed recordings where the method by Amundsen (2001) has been applied with some success (Brunellière et al., 2004).

In seabed seismic data acquisition where multicomponent geophone and hydrophone recordings are commonly made, it is possible to decompose the recorded data into both up- and downgoing components as they would be perceived just above the seafloor (in the water) or just below inside the seabed (Amundsen, 1993b). If the data are decomposed into their up- and downgoing components in the seabed (below the water layer), the receiver-side water-layer peg legs will automatically be excluded from the upgoing part of the wavefield. This has often proved to be a very effective means for removing significant multiple energy in seabed recordings (Stewart et al., 2007).

The level of the sea surface is affected by, for instance, tides altering the length of the propagation leg of surface-related multiples in the water. This aspect was exploited by Calvert (2005a), Robertsson and Kostov (2010), and Ikelle (2004) to develop multiple attenuation methods using seismic data acquired in two different tidal states.

Finally, multiples do not necessarily have to be regarded as noise. Provided that the multiples can be processed properly, they convey information about subsurface structures and can even help to image structures that have not been illuminated by primary reflections (e.g., when undershooting surface obstructions such as oil platforms or in the vicinity of complex salt structures). This aspect has been highlighted recently in a similar context to simultaneous source acquisition (see **Section 11.06.5.2.2**) where it has been proposed to fundamentally change processing of seismic data moving away from the view of acquisition and separation of sequential shot records. Instead, an alternative approach would be to consider illumination of the subsurface achieved by any means of acoustic energy radiated from different locations on the surface, be it from simultaneous sources or surface-related multiple reflections (Berkhout and Verschuur, 2012; Muijs et al., 2007a,b; Wapenaar et al., 2012).

11.06.3.7 Seabed Seismic-Specific Issues

11.06.3.7.1 Scholte waves

Just as in a land seismic survey where Rayleigh waves (ground roll) are a major source of noise, surface waves are sometimes also generated at the acoustic–elastic seafloor interface. These correspond to Scholte waves and are often referred to as mud roll. Mud roll is generated in shallow water when the emitted seismic energy couples into the surface wave modes. Although it can sometimes (but not often) be observed as an evanescent arrival on towed marine seismic data, it is more commonly observed on seabed seismic recordings. Due to their slow propagation velocity compared to primary reflections, mud roll is relatively straightforward to filter from the recorded data. As opposed to ground roll on land where near-surface properties can undergo significant lateral variations, mud roll tends to suffer less scattering and propagates more uniformly.

Scholte waves can also provide important information about the elastic properties of the shallow seabed. For example, Olofsson et al. (2003) inverted Scholte wave data for seabed velocities in the Valhall field in the North Sea where anisotropic velocities were consistent with subsidence of a production platform.

11.06.3.7.2 Vector fidelity

In seabed seismic recordings, the geophone records the particle velocity of propagating waves in the seabed. Whereas the vertical component of particle velocity is continuous across the fluid–solid interface, the horizontal components are not. The faithful recording of the particle motion is referred to as vector fidelity (Dellinger and Clarke, 2004; Gaiser, 2003). However, coupling of the geophones to the seabed and resonances in the seabed cable acquisition system may lead to severe distortions of the recorded wave motion. Therefore, considerable efforts have gone into the design of seabed seismic acquisition systems to enable a high degree of vector fidelity. One example is the Q-Seabed® system developed by WesternGeco where 'boots' were positioned around sensor housings to prevent them from a rolling motion induced by horizontally polarized shear wave arrivals from the subsurface (see **Figure 8**; Kragh et al., 2004). Different processing methods have also been developed to compensate for insufficient vector fidelity in the recording system (e.g., Schalkwijk et al., 2003).

11.06.4 Acquisition Systems

In marine seismic exploration, quantities relating to the wavefield are measured. These are typically the wavefield pressure in the water and perhaps particle motion (acceleration or velocity) as well. The measurements are made over the seismic bandwidths (up to a few hundred hertz) and with a dynamic range that can reach 140 dB. In addition to measurements of the seismic wavefield, there can be measurements of other quantities such as the local magnetic field, data from acoustic positioning systems, data relating to the monitoring of marine mammals, bathymetry, weather conditions, and ocean salinity and temperature. It may also be necessary to specifically measure the signal and noise that propagate within the structure of the streamer itself so that the noise can be attenuated. In conventional towed marine systems, these data are transmitted to the vessel at once. In seabed and autonomous systems, they can be either stored near the sensor for later recovery or transmitted to a recording vessel or platform.

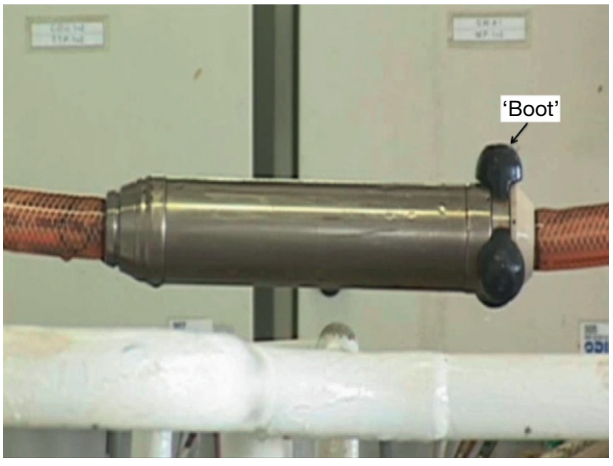


Figure 8 A seabed seismic sensor (cable type). It rests on the bottom and measures particle motion of the Earth and pressure in the water. Photograph provided courtesy of WesternGeco.

11.06.4.1 Marine Recording Systems

In the 1980s and earlier, it was standard practice that the seismic hydrophones in a streamer were hardwired into groups and a twisted pair of copper wires would convey the analog data from each group up to the vessel. The data would be digitized and multiplexed on board to form a data stream, which was stored on tape. Because of the multiplexing, the data samples were not read simultaneously but with a small progressive delay from one receiver group to the next. This 'skew' could be corrected in processing. The data were typically stored in a floating-point notation but where the exponent followed a predetermined time curve and which was the same for all the channels. In the most modern systems, the digitization takes place at the sensor and the data are sent to the vessel along high-bandwidth fiber-optic pathways using standard floating-point digital technology. In some systems, the individual hydrophones are separately recorded; in others, they are hardwired into groups within the streamer and only the summed output is transmitted to the vessel.

Seismic streamers contain four main components: stress members, sensors, filling, and an outer skin. Stress members (steel or Kevlar ropes typically) carry the overall tension, and the sensor packages are loosely anchored to them. Circular spacers are placed at intervals and these hold the outer skin (usually polyurethane) in position. The interstitial spaces used to be filled with kerosene, and the quantity could be adjusted to make the streamer neutrally buoyant for the particular survey conditions. However, kerosene fell out of favor because of the risk of fire and of spills into the ocean; modern streamers are either solid, foam, or gel. Steamers are manufactured in sections each typically 100 m long with connectors at each end. Figure 9 shows some kerosene-filled streamers stored on drums.

Figure 9 Seismic streamers coiled on their drums for storage. The transparent polyurethane skin can be seen and inside it the yellow spacers and the bundle of cables wrapped around the stress members. The filling is (transparent) kerosene. The sensor packages are located between each close pair of spacers. Photograph provided courtesy of WesternGeco.

The leading ends of the streamers are, of course, attached to the vessel, and in cases of multiple streamers, there will be a large deflector (hydrofoil) to move each streamer to the correct lateral position (the lower panel of Figure 1 shows some detail of the front-end architecture of the spread). The tail of the streamer is linked to a tailbuoy on the surface, which provides tension and incorporates a GPS positioning system. In a large 3D array, there might be 12 or more streamers towed behind the vessel. The towing and handling issues with all this equipment are considerable. Risks to the streamers are tangling, breaking under tension, barnacle growth, and, occasionally, shark attack.

When hardwired groups are used, the signal-to-noise ratio is improved as long as the separation between hydrophones is larger than the coherence length of the noise. This is normally the case for the turbulence noise because it is generated locally. In addition, some noise can propagate within the streamer with a low velocity and a correspondingly high wave number. The overall length of a group affects the inline wave number response of the system: sound waves traveling perpendicular to the streamer will be in phase for all hydrophones and will add constructively; waves that are arriving at an angle to the streamer will experience a phase shift along the group, and this will result in attenuation for wavelengths that are not large compared with the group length. The result is a spatial 'wave number' filter. In many systems, the groups overlap and the weighting is nonuniform. This is done to avoid a ringing response in the wave number domain that would otherwise be associated with a steep-sided spatial aperture. In more modern systems, the hydrophones are recorded separately and there are no hardwired groups. The noise field is fully sampled by the closely spaced sensors, and advanced digital filtering can be used to remove it rather than simple summation.

When individual sensors are recorded (not hardwired into groups), filters can be designed to remove noise more effectively than is possible with hardwired groups (Özbek, 2000). By using single-sensor recording and advanced filtering, it is possible to remove the noise so well that data can be acquired while the vessel is turning and the streamers are in cross flow, which leads to higher noise in the individual hydrophone signals. This, in turn, enables data to be acquired using a moving circular path ('coil shooting'; see **Section 11.06.5.1** and Brice et al., 2012) such that no time is lost in changing from one line to the next (Moldoveanu, 2008; Taylor et al., 2009).

In addition to measuring the wavefield pressure, some streamers today measure the vertical particle motion as well. Barr and Sanders (1989) showed the value of vertical motion measurements for deghosting. Söllner et al. (2007) and Tenghamn et al. (2007) showed the results from GeoStreamer® (from Petroleum Geo-Services), which uses hydrophones and gimbaled (vertical component) geophones. Robertsson et al. (2008) and Özbek et al. (2010) discussed the uses of a multi-measurement streamer (measuring horizontal and vertical particle motion and pressure). In 2012, such streamers became a commercial reality with Nessie-6® (from WesternGeco), which uses hydrophones and microelectromechanical accelerometers. The inline component of particle motion is related to the inline pressure gradient, and since the pressure measurement is well-sampled, the inline particle motion measurement is, in principle, not needed.

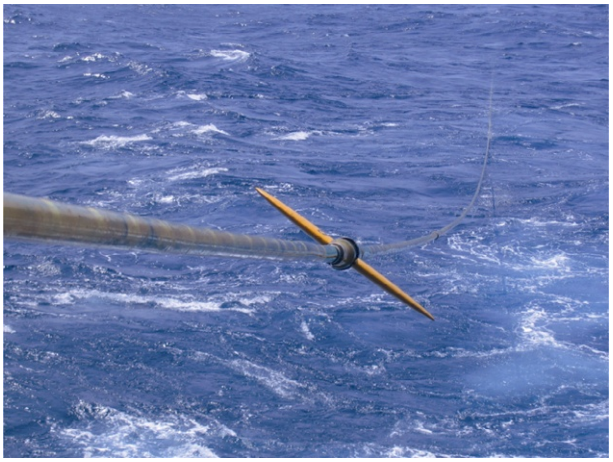


Figure 10 A control fin on a streamer during deployment of the streamer. Photograph provided courtesy of WesternGeco.

The depth of a modern streamer is controlled by fins using hydrostatic pressure for depth measurement. Some fin designs allow the streamer to be steered laterally as well (Bittleston et al., 2000). This lateral steering allows the receivers to be placed closer to their predetermined positions – a great advantage for time-lapse studies – and has operational and safety advantages also. In order to make use of lateral steering, there needs to be an accurate positioning system along the whole length of the streamer array. This can be achieved using a fully braced acoustic network linked to GPS receivers. An example of a streamer control fin is shown in **Figure 10**.

As discussed in **Section 11.06.3**, the noise recorded by the sensors in a streamer originates from several sources. A major part comes from turbulent eddy shedding generated by the streamer and has a coherence length that can be small compared with the acoustic wavelength. Recently, some research has been reported on methods for reducing the turbulence noise by special coatings (Elboth et al., 2010; Martell et al., 2010). Another major source of noise is the longitudinal vibration from the towing vessel and lead-in cables. This wave propagates along the stress members at speeds that are significantly faster than the speed of sound in water (6000 m s^{-1} is typical). There is also noise that is present in the survey area independent of the seismic survey, such as shipping noise. The noise, whatever its source, also propagates within the streamer. In fluid-filled streamers, the mode of propagation is the 'bulge wave.' In this wave, the kinetic energy of the fluid filling is coupled to the elastic potential energy of the dilated streamer skin. Such bulge waves propagate slowly (and are dispersive); 40 m s^{-1} is typical. In solid- and gel-filled streamers, there can be a number of more elaborate propagation modes within the streamer. In general, the noise becomes worse when the weather is bad and when the streamer is in cross flow, such as when it is towed in a curve.

11.06.4.2 Seabed Recording Systems

The evolution of seabed recording systems has been varied, but in general, such systems aim to measure the pressure in the water and the three components of particle motion of the solid seabed, for example, Berg et al. (1995) described such a sensor. De Kok (2012) is a good general reference. Some sensors also measure the seabed slope with a tilt meter (Pettenati-Auziere et al., 1997). Such systems are valuable for time-lapse seismic because of their repeatability and for other purposes because of the multicomponent nature of the acquisition. For example, they directly measure shear waves in the subsurface. They also allow a wide range of azimuths (Beaudoin and Ross, 2007) to be simultaneously recorded from a single shot. The first commercial survey done by the oil and gas industry with seabed receivers was at the Tommeliten field in the North Sea in 1993 (Berg et al., 1995).

Some seabed sensors are planted into the subsurface much like land geophones (Berg et al., 1995), whereas others merely stand on the seabed (Pettenati-Auziere et al., 1997) being held on the bottom by their own weight. Others formed part of a cable that could be dragged into opposition and possibly buried (Cooper et al., 1999). In early (academic) crustal-type surveys, ocean-bottom seismometers were weighted and dropped from a surface vessel and later recovered after they had released themselves from their weights after the survey (Kirk et al., 1982). This practice remains common today.

Among others, Amundsen et al. (2001) showed how seabed systems could be used for multiple attenuation, and Martin and Davis (1987) and Bale et al. (2000) reported the use of seabed data for detecting birefringence of shear waves. Veire and Landrø (2006) discussed joint inversion of PP and PS wave data.

There are considerable logistic difficulties in installing a seabed sensor (Berg et al., 2008). Many different designs have been tried in order to improve the coupling (Berg et al., 2002; Woje et al., 2002). The scattered shear wave energy received by the sensors means that their positions must be accurately known (Hays et al., 2008). **Figure 8** shows an example of a seabed sensor node. The data from seabed sensors either can be sent directly to the surface or can be stored at the sensor for later collection. In the oil and gas industry, seabed systems are more typically used over existing producing reservoirs where surface obstructions sometimes prevent towed marine streamer operations or they are used to monitor the depletion and perform time-lapse studies rather than for 'frontier' imaging. An example of integrating seabed systems with towed marine systems is given by White et al. (2002) for complex imaging below basalt over long survey lines.

Usually, the seismic source is deployed in the water in a conventional way (either from a vessel or, in the case of VSP – vertical seismic profile – from a rig) although seabed sources have also been devised (e.g., Socco et al., 2011).

11.06.4.3 Marine Sources: Air Guns

By far, the most commonly used marine seismic source today is the marine seismic air gun. Parkes and Hatton (1986) is a good overall and detailed reference. Air guns were developed in the 1960s at Lamont Doherty Earth Observatory, and Ewing and Tirey (1961) first described them; the primary driver was to find a safer source to replace dynamite. The idea was developed into a viable commercial seismic source by Bolt Technology Inc. Other air guns were produced by Prakla, Geophysical Services Inc. (the sleeve air gun), and Sodera (the 'G-Gun'). A photograph of a cutaway air gun is shown in

Figure 11. Other sources, such as dynamite (Maxipulse®), steam-generated cavities (Vapochoc®), spark-induced cavities ('sparkers'), and momentum-induced cavities (water guns), have been used but suffer from various drawbacks. Marine vibratory sources have been used in the past and may experience resurgence because of possible environmental and geophysical advantages.

An air gun is essentially a container of compressed air with a vent. When commanded to do so, the vent opens and rapidly releases the air into the water. This release and the subsequent oscillation of the air bubble cause sound to be radiated into the water. The mechanism by which the vent is opened varies between air gun types, but they all work essentially by using the air pressure to operate the mechanism. The air presses on two opposing surfaces of the moving part of the vent. The areas of these two surfaces are slightly different, and the resulting force imbalance is just sufficient to hold the vent closed. A small extra air path, opened by a solenoid, upsets that force balance, and the vent opens rapidly under the full force of the air pressure. The air bubble expands under pressure until it is overcome by the hydrostatic head of sea water, which then causes an implosion. The process repeats as the bubble rises under buoyancy such that the 'source' comprises a series of expanding and collapsing bubbles. When the bubble finally breaks the surface, there are no more oscillations.

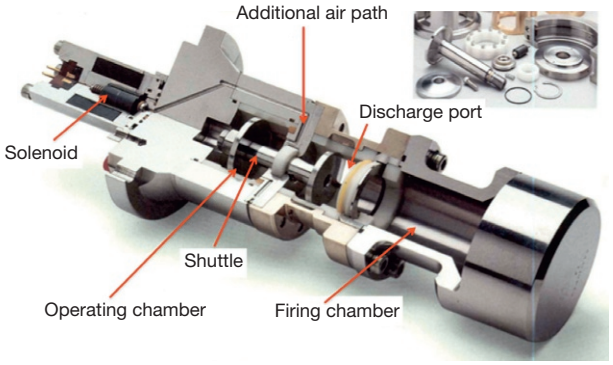


Figure 11 A Bolt air gun. Courtesy of Bolt Technology Inc.

Figure 12 shows a sequence of photographs of the released bubble. The bubble pulsates in the water and radiates pressure peaks at the radius minima. The air gun firing mechanism is very elegant; it has only two moving parts, and one of those is the tiny solenoid actuator. The sudden release of compressed air causes a pressure pulse with a relatively steep rising edge followed by a damped oscillation. The bubble is small compared to the wavelength of the sound it radiates, so it behaves as a monopole at distances that are large compared with its size. Its sound field is characterized by a single monopole source function, which is related simply to the volume acceleration of the bubble. Another consequence of the smallness (compared with wavelength) of the bubble size is that the acoustic radiation is inefficient; much energy flows back and forth between compressional energy of the air inside the bubble and kinetic energy in the surrounding water, but little of it is radiated as sound. The efficiency is improved when they are used in an array (Laws et al., 1988).

For any small (compared with wavelength) acoustic radiator, the monopole source function is related to the volume acceleration by eqn [1] (see, e.g., Kinsler et al., 1982):

$$S(t) = \frac{\rho}{4\pi}\frac{\mathrm{d}^2 V(t)}{\mathrm{d}t^2} \qquad [1]$$

where $V(t)$ is the volume of the monopole, ρ is the density of water, and $S(t)$ is the notional source signature. The pressure field of a monopole is given by

$$p(r,t) = \frac{1}{r}S\left(t - \frac{r}{c}\right) \qquad [2]$$

where $p(r,t)$ is the pressure measured at a distance r from the monopole and c is the propagation velocity of the sound wave

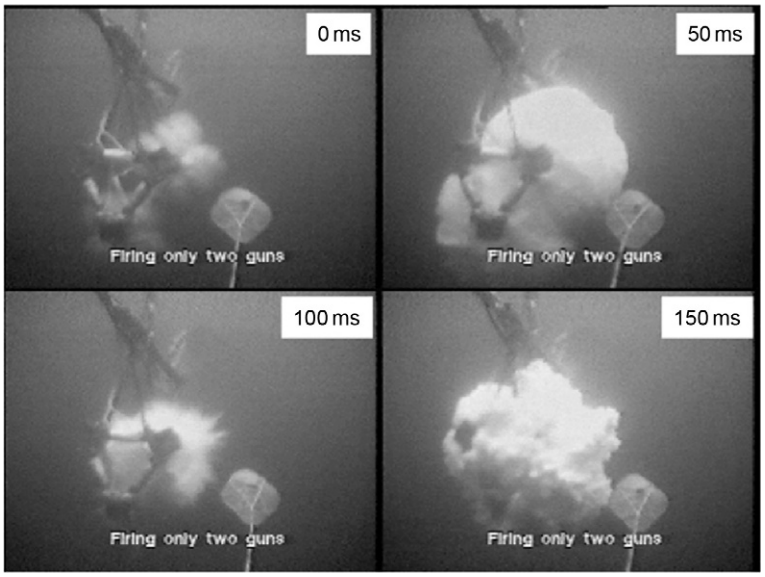


Figure 12 A pair of Bolt 150 cubic inch air guns photographed at about 50 ms intervals. Just after firing, at peak radiated pressure (top left); at first maximum radius (top right); at first collapse (bottom left); at second maximum radius (bottom right). The air guns are about 1 m apart. Photographs provided courtesy of WesternGeco.

in water. Thus, for a pulsating source that is small compared with the wavelength, we have

$$p(r,t) = \frac{1}{r}\frac{\rho}{4\pi}\left(\frac{\mathrm{d}^2 V\left(t-\frac{r}{c}\right)}{\mathrm{d}t^2}\right) \quad [3]$$

Air guns are used in arrays, often of several tens of units. This is done for three principal reasons: (1) to increase the overall acoustic output, (2) to obtain a suitable spectrum, and (3) to obtain a desirable directivity pattern. The deployment of air guns in an array as distinct from the use of clusters (air guns at very small separations similar to the maximum bubble radius) is discussed elsewhere (e.g., Dragoset, 1990).

A single air gun bubble emits a spectrum with a pronounced main peak and harmonics. In order to obtain a flatter spectrum, bubbles of different volumes are released together so that their different resonant frequencies combine to give a smoother spectrum. This is illustrated in Figure 13. Seven air guns are used with a range of volumes, and the individual notional source signatures are shown in the upper panel. The spectra are shown in the lower panel. The bubbles are placed sufficiently far apart that the coupling between them is small and they each oscillate at their own characteristic frequency.

The pulse that is emitted in the vertical direction (the 'vertical far-field signature') is the sum of these pulses added to the sum of the pulses reflected by the sea surface. This is illustrated in Figure 14 where the effect of using three different deployment depths is shown. Changing the deployment depth has two effects: (1) it alters the hydrostatic pressure and thereby the resonant frequency of the bubbles and (2) it alters the time delay between the direct and the surface-reflected (the 'ghost') signal. At very low frequencies (below the lowest bubble resonant frequency), these two effects almost exactly cancel (Hopperstad et al., 2012).

An array of air guns is not small compared to the wavelength, so the geometric layout has a significant effect on the directivity of the source. In addition, the presence of the sea surface, with its pressure reflection coefficient of -1, adds a 'ghost' image of the source array at a height above the sea surface equal to the source depth and tends to make the directivity primarily that of a vertically oriented dipole. Directivity is illustrated in Figure 15 for the vertical and nonvertical (45°) directions. In blue are shown the contributions of the seven (in this case) air guns in the array, and in red is shown the combined far-field pulse. In the vertical direction, the main peaks are synchronized and the total pulse amplitude is high. In other directions, time delays cause the main peak to be reduced.

When 2D seismic profiles are being surveyed, it is usual to design the source so that energy is emitted mainly in the plane of the survey. This means that wide arrays are used. For 3D surveys, a monopole source would be ideal and small arrays are preferred. This is discussed further in Section 11.06.5.2.3.

It is not necessary that all the air guns be deployed at the same depth, although this is often what is done. If all the air guns are at depth d, then there will be a zero (the 'ghost notch') in the vertical far-field spectrum at a wavelength of $2d$. This is due to the cancelation of the direct signal with the phase reversed and delayed surface-reflected signal. Typically, the air guns are deployed at a depth of about 6 m so the ghost notch is at about 125 Hz.

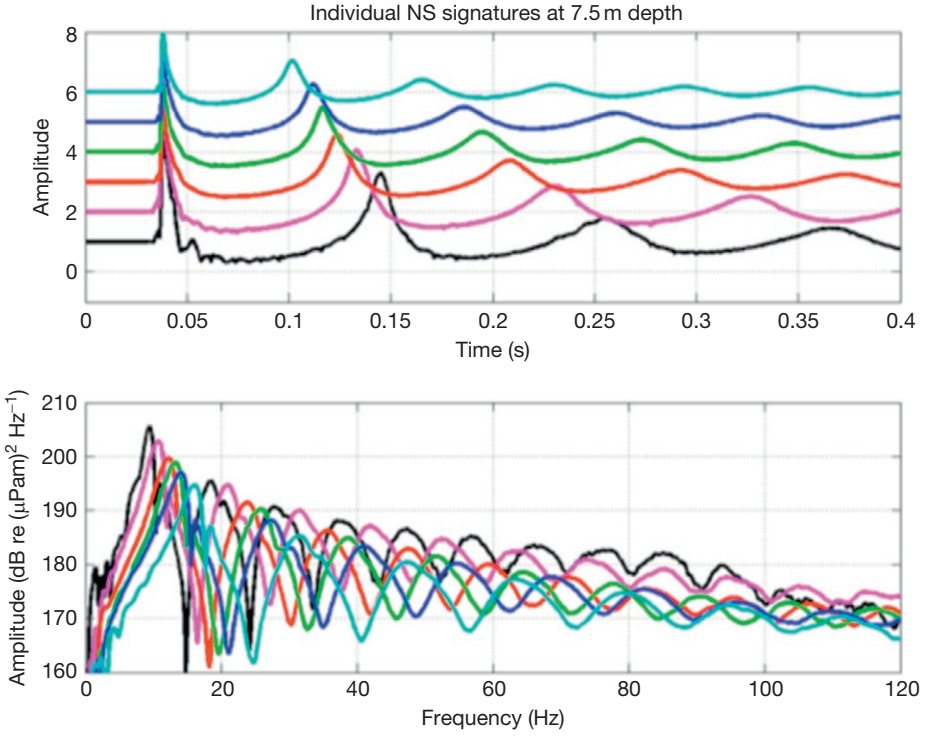


Figure 13 A mix of air gun volumes is used so that there is a range of oscillation periods in the combined output. This has the effect of flattening the spectrum. The individual notional source signatures are shown in the upper plot and the energy spectra in the lower plot.

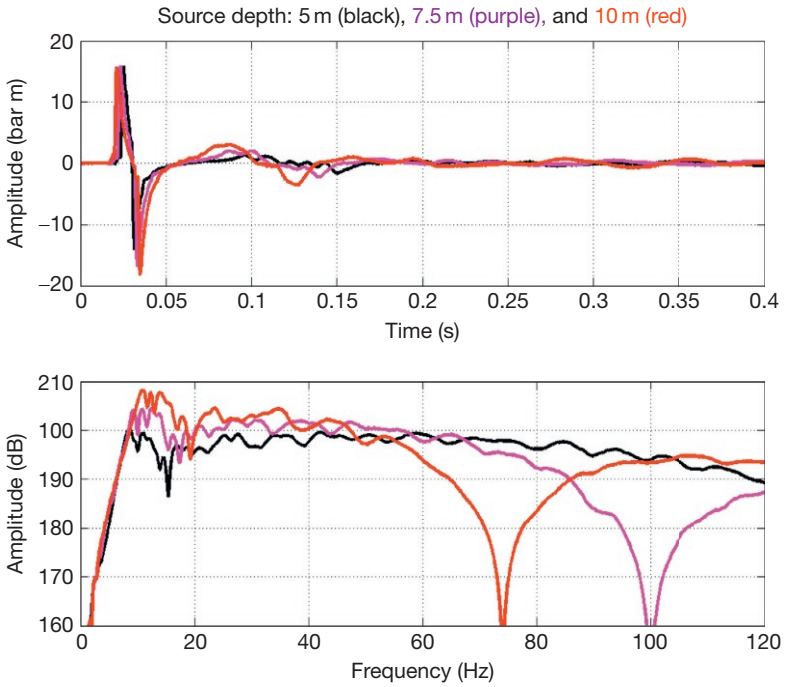


Figure 14 Combined vertically traveling far-field signature of the array. Notice the sea-surface reflection (the ghost) with its negative peak following shortly after the main positive peak. This results in a notch in the spectrum whose frequency depends on the source depth.

The roughness of the sea surface affects the source signature in two ways: both (i) because the source is buoyed from the surface (usually) and therefore moves up and down with the waves and (ii) the sea surface is not flat and so the scattered (ghost) wavefield is modified. Laws and Kragh (2002) showed that this effect (together with a similar effect on the receivers) is significant in time-lapse seismic if it is left uncorrected. They also showed how to correct it (Laws and Kragh, 2006, 2009).

When VSP surveys are conducted, it is usual to employ a smaller number of air guns than is usually available for a large towed marine survey. In such cases, clustered air guns are typically used (Dragoset, 1990). Although such sources emit a less flat spectrum, the clustering does lead to a more damped oscillation. The VSP source obtains its spectral flatness by using clusters rather than by using a large mix of volumes. The downgoing pressure sound pulse that has been reflected in the sea surface begins with a strong negative peak, and in some places, in the water near the middle of the array, this can overcome the ambient pressure and cause cavitation. Landrø et al. (2011) had reported this with observations of the sound made by the cavity collapse.

Ziolkowski et al. (1982) and Parkes et al. (1984) showed how the monopole source function (termed the 'notional source signature' by them) could be measured using a hydrophone placed about 1 m from the air gun bubble. Crucially, they showed how this could be done even when an entire array of air guns were being fired together by using a set of near-field hydrophones with each one placed close to one air gun in the array. The positions of the air guns and their notional sources completely define the output field of the array: the pulse shape can be computed in any direction.

The notional source representation of Ziolkowski et al. (1982) was also used in other proposed methods for measuring the source signature. Amundsen (1993a) and Landrø and Sollie (1992) both proposed using data from a small streamer towed under the source array to estimate the notional sources. Landrø and Sollie (1992) constrained them to be solutions of the bubble equation of motion (described in the succeeding text); Amundsen (1993a) allowed them to be arbitrary functions. These two methods and that of Ziolkowski et al. (1982) combined with Parkes et al. (1984) were compared by Laws et al. (1998) to measurements made with a far-field hydrophone, and the three methods were found to be of similar accuracy. Ziolkowski (1980) had proposed a method using 'scaled arrays,' but this involved doubling the shot interval and was not adopted. Only the notional source method is in current use today. In the absence of a direct measurement, the source signature must be estimated from the data themselves, perhaps implicitly using predictive deconvolution (Ziolkowski, 1984). Estimating the signature from the recorded seismic data implies that the frequency spectrum of the underlying seismic data is white (but band-limited) and that the signature is minimum phase. Although neither of these assumptions is strictly correct (Ziolkowski, 1984), the method is often accepted to be sufficient in practice.

Work on the equation of motion of pulsating bubbles and cavities in fluids has been ongoing since Rayleigh (1917) published his paper on the collapse of a cavity in a fluid. Although cavities and bubbles behave in different ways, there are strong similarities in the equations that govern their motion. There is also a strong similarity in the radial motion of spherical bubbles ranging in size from very small ultrasonic bubbles to very large explosion bubbles. Much early pioneering work was done on explosion bubbles by such authors as Herring (1941), Kirkwood and Bethe (1942), Plesset and Zwick (1952), Gilmore (1952), and Keller and Kolodner (1956).

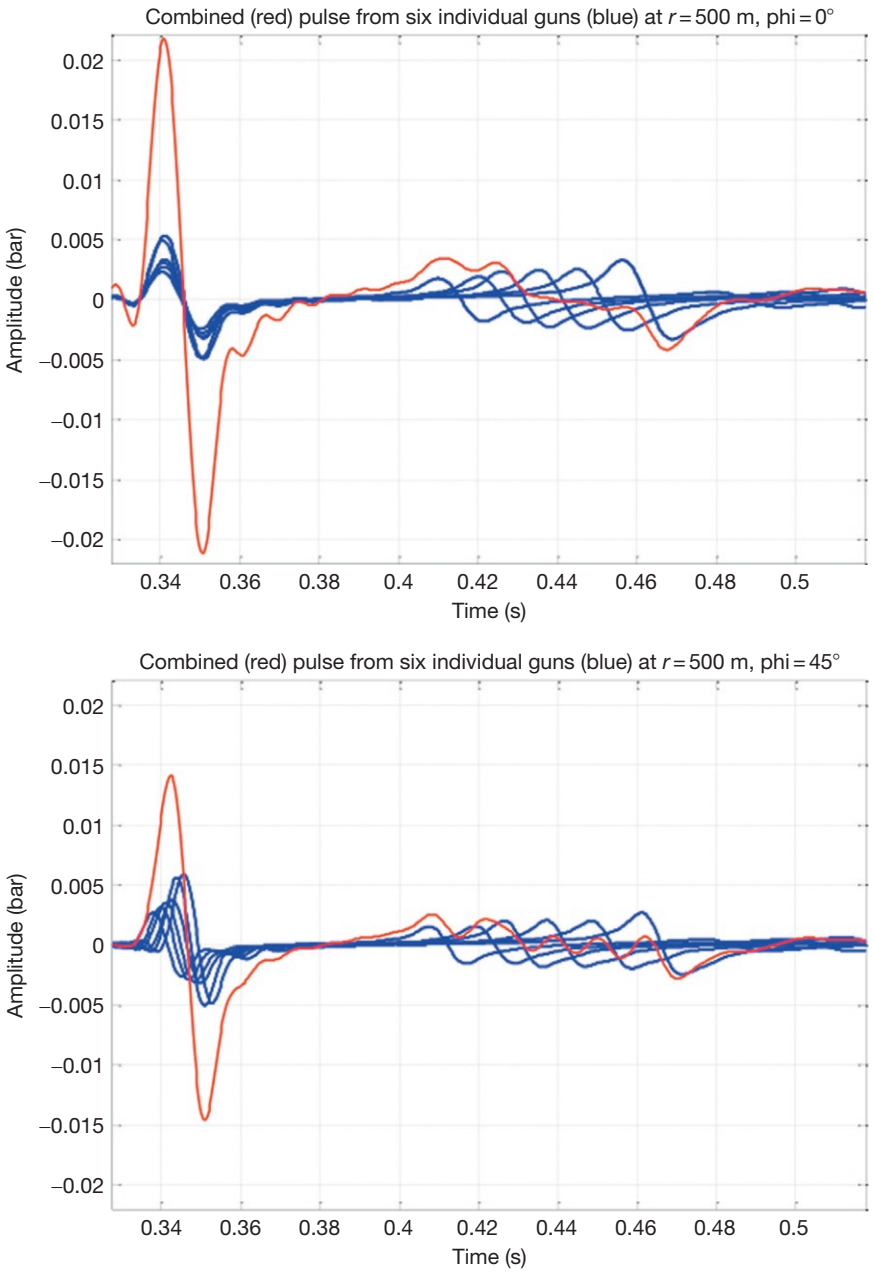


Figure 15 The far-field signature at a distance of 500 m in a vertical direction (upper panel) and at 45° (lower panel). The six pulses from the six air guns in the array are shown in blue and their sum is shown in red. Note that the constructive interference in the vertical direction gives a higher pulse amplitude than at 45° to the vertical.

The system is essentially a spring–mass oscillator with the air in the bubble forming the spring and the surrounding water forming the mass.

In 1970, Ziolkowski introduced air gun bubble modeling to the geophysics community. He based his work on the Gilmore (1952) equation. Schulze-Gatterman (1972) used the formulation of Keller and Kolodner (1956). Air gun bubbles in an array interact with each other because they each oscillate in a time-varying pressure field. This caused complications when designing an air gun array, and Giles and Johnston (1973) published criteria for the separations at which the interaction could be neglected. Vaage et al. (1983) showed experimental results on how the signature of a single air gun was related to the firing pressure, the volume, and the depth. They also compared their results with formulas derived from various other papers. In 1984, Vaage et al. (1984) published a paper showing experimentally the effects of the interaction of air guns, and Strandenes and Vaage (1992) measured the effect of air gun clustering (separation similar to maximum bubble radius). Ziolkowski and Metselaar (1984) showed how to model the interaction between bubbles using the same approach that Fujikawa

and Takahira (1986) used with ultrasonic bubbles. Laws et al. (1990) showed how clustered air guns could be modeled including the effects of frequency locking when the bubbles become very close. Laws et al. (1988) showed experimentally that the energy emitted by an array of seven air guns fired simultaneously was about double the amount emitted if the same air guns were fired sequentially. This indicates that the efficiency was different in the two cases because the input energy (to charge the air guns) was the same in both cases. They also showed that there would be similar effects for marine vibrators.

The Rayleigh–Plesset equation for a spherical bubble in a liquid is given by eqn [4]. Rayleigh (1917) derived it for describing the collapse of a spherical cavity and Plesset (1949) applied it to cavitation bubbles:

$$\frac{P_b(t) - p_\infty}{\rho} = R\frac{d^2R}{dt^2} + \frac{3}{2}\left(\frac{dR}{dt}\right)^2 + \frac{4\upsilon}{R}\frac{dR}{dt} + \frac{2T}{\rho R} \qquad [4]$$

where $P_b(t)$ is the pressure within the bubble, assumed to be uniform; p_∞ is the local ambient pressure; ρ is the density of the water; $R(t)$ is the radius of the (spherical) bubble; υ is the kinematic viscosity of the surrounding liquid, assumed to be constant; and T is the surface tension of the bubble.

The equation of state for the air inside the bubble must be supplied, and since, in general, pressure is not a function simply of volume (but also of temperature), some assumptions about the heat transfer must be made. The viscous and surface tension terms are negligible for air gun bubbles, and the radiation damping is also negligible, which is because the bubble is small compared with the wavelength of the sound it radiates. The bubble is heavily damped so eqn [4] does not tell the whole story.

The air gun bubble radius is typically about 0.5 m. If two air guns are fired simultaneously, there are various degrees to which the two oscillations are coupled. When the bubbles are many radii apart (say 10 m), there is almost no interaction and the emitted pulse shapes are the same as when the bubbles are fired independently. If the air gun volumes are different, then the bubbles will oscillate with different periods. At about 3 m, the oscillations are still independent but there is some coupling. At about 1.6 m, the oscillation periods 'lock' and the bubbles oscillate in phase and at a frequency that is close to the resonant frequency of a single bubble with a volume equal to the sum of the volumes of the bubbles. At separations of about 1 m, the damping effects are greatly increased (clustering). These effects are discussed, for example, in Vaage et al. (1984), Laws et al. (1990), and Hopperstad et al. (2012).

For a time, it was popular to insert a 'wave shape kit' into an air gun to subdivide the air chamber into two interconnected spaces, but this is no longer normally done. It increased the damping but did not affect the period (Vaage et al., 1983).

The interaction between bubbles is usually included by modifying the ambient pressure p_∞ to include the pressure field from the other bubbles and the reflected signal from the sea surface (and possibly the seabed as well; Ziolkowski and Metselaar, 1984). Further semiheuristic corrections (but using the spherical bubble approximation) to the interaction can be added to account for bubbles that are within a few bubble radii (Laws, 1991; Laws et al., 1990). Cox (2003) and Cox et al. (2004) gave a more thorough account of the interaction by accounting for nonspherical bubble effects.

The air pressure in the bubble depends on the thermodynamic state of the air, and this is not a function of the radius alone. Some assumptions about the heat transfer and evaporation are needed in order to solve this equation. Ziolkowski (1970) proposed using an adiabatic equation of state (making the pressure a function of radius) but with a modified polytropic exponent. He also applied a heuristic decay to the radius in order to damp the motion. Laws (1991) included evaporation and condensation at the air/water interface, the thermodynamic behavior of the fog inside the bubble, and heat transfer across the interface. He showed that the water vapor in the bubble had negligible effect, that radiation damping was insignificant, but that the heat transfer could explain the damping if the heat-transfer rates were increased (possibly justified by the turbulence). However, the damping was eventually shown by Cox et al. (2004) to result from turbulence and jetting both of which are absent in the spherically symmetric models. The jet forms on account of the ambient pressure gradient in the water, which is itself a result of the hydrostatic pressure gradient and from the pressure field of any nearby bubble.

Another development has been the GI® 'generator–injector' air gun developed by Sodera (Pascouet, 1991) and currently marketed by Sercel. This is essentially two air guns ('G-guns') that are fired sequentially. The second gun is fired when the first bubble reaches maximum radius, and the air released by the second gun brings the air pressure inside the bubble up to ambient. Since, at that point, there are no kinetic energy in the water and, as a result of the step in pressure, no potential energy in the air, the oscillation stops abruptly. Landrø (1992) gave a more complete description and also discussed how to model such an air gun.

Prosperetti (1984) reviewed the physical aspects of the behavior of bubbles in sound fields. He showed that all the existing spherical formulations are correct to first order in Mach number and differ only in the selection of the second-order terms they include.

The well-known Rayleigh–Willis equation relating the bubble period to the volume, pressure (in fact to the mass of air), and depth of the bubble was derived by Willis (1941) using an energy argument:

$$f \propto \left(1 + \frac{d}{d_h}\right)^{5/6} (pV)^{1/3} \qquad [5]$$

where d is the depth of the bubble, d_h is atmospheric pressure expressed as a head of water (10 m), p is the initial bubble pressure, and $V$is the initial bubble volume.

When air guns are fired from cold, they gradually warm up over repeated shots and the changing temperature affects the pulse shape (Langhammer and Landrø, 1993; Parrish, 1985, 1986). The bubble periods reduce and settle to a smaller value over the first several shots. Perhaps, the Rayleigh–Willis equation might better be couched in terms of the mass of air in the bubble rather than the product pV.

The environmental impact of seismic air guns has come under increasing scrutiny; Richardson et al. (1995) is an excellent book on the subject. Various measures are employed

to mitigate potential damage to marine life, especially to marine mammals. One measure is the use of an 'exclusion zone' around the seismic source; the source is stopped if a marine mammal is found within the exclusion zone. A second measure is 'ramp up' whereby the strength of the source is gradually increased over a 20 min period at the start of each survey line. Southall et al. (2007) collected together all the current information on marine mammal hearing damage and defined acoustic criteria for assessing the likelihood of injury. Laws (2010) applied these criteria to measured air gun pulses and showed that the zone of possible injury was smaller than the typical exclusion zone around the seismic source (**Figure 16**). Breitzke and Bohlen (2010) obtained similar results. The effectiveness of ramp up is unclear. The Southall et al. (2007) criteria indicate that marine vibrators would carry no risk of hearing injury to marine mammals; the injury threshold is higher for nonimpulsive sources. In addition, Richardson (2011) concluded that vibrators would be less disturbing (as opposed to injurious) to marine mammals than air guns are. Although definitive data on disturbance are not readily available, several studies are currently under way. It has been suggested by Laws et al. (2008) and Kragh et al. (2012) that seismic sources may sometimes be unnecessarily loud, particularly at the higher frequencies. An air gun with reduced high-frequency output is currently under development (Hopperstad et al., 2005); the prototype version is shown in **Figure 17**. In addition, the high-frequency output can be reduced by desynchronizing the firing times (Hopperstad et al., 2008b). It has also been proposed to use curtains of bubbles to attenuate the lateral propagation of energy from a seismic source (Behrens, 1999). Of course, with marine vibrators, the frequency content is more controlled and can, in principle, be confined to the frequency range needed for imaging. This advantage is in addition to the advantage of the higher damage threshold that is used for nonimpulsive sources given by Southall et al. (2007).

11.06.4.4 Marine Sources: Vibrators

Marine vibrators have been used in the past and there is growing interest in them (Richardson, 2011). There are several possible designs, but they all work by creating a volumetric displacement of the surrounding water. For a vibrator that is small compared with the wavelength it emits, the notional source signature is proportional to the volume acceleration, so a limitation in volume displacement leads directly to a low-frequency roll-off of the output. This is the main reason why vibrators have not been successful in the past; they lack low-frequency power. This limitation is shown in the equations in the succeeding text.

Let us calculate the radiated acoustic power from a vibrator operating at frequency f with a swept volume satisfying

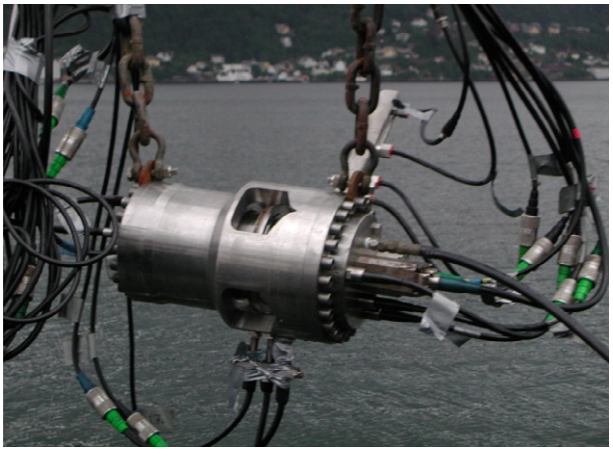


Figure 17 The E-source: a prototype air gun with a reduced high-frequency output being developed jointly by Schlumberger and Bolt Technology Inc. Photograph provided courtesy of WesternGeco.

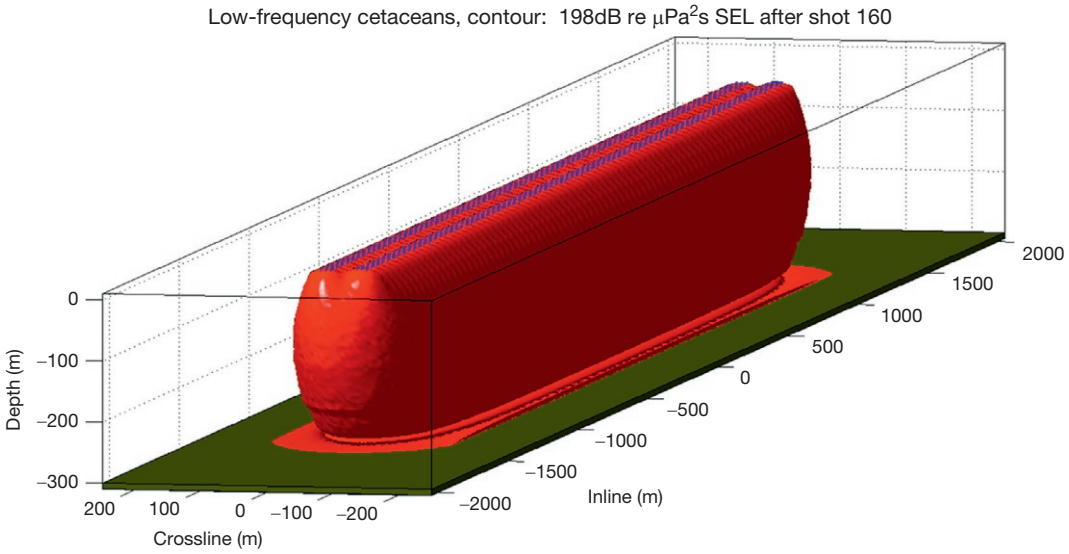


Figure 16 Potential injury zone around a seismic line following the Southall et al. (2007) criteria for impact of low-frequency acoustic emissions on marine mammals (from Laws, 2012). The small blue marks represent the seismic shot positions. A stationary low-frequency cetacean within the red zone would risk injury. However, in many jurisdictions, the mitigation exclusion zone shuts down the air guns before such a situation arises.

$$V(t) = V_0 \sin(2\pi f t) \quad [6]$$

Recall that, for a small pulsating source,

$$p(r,t) = \frac{1}{r}\frac{\rho}{4\pi}\left(\frac{\mathrm{d}^2 V}{\mathrm{d}t^2}\right)_{t-(r/c)} \quad [7]$$

so

$$p(r,t) = -\frac{1}{r}\frac{\rho}{4\pi}(2\pi f)^2 . V_0 . \sin\left(2\pi f . \left(t - \frac{r}{c}\right)\right) \quad [8]$$

The mean square monochromatic power output (intensity, measured in Watts) is given by integrating the power over the surface of a sphere of arbitrary radius, r, and averaging over time:

$$I(f) = \frac{1}{\rho c}\int_{4\pi r^2} \langle p^2(r,t)\rangle \mathrm{d}A \quad [9]$$

so that

$$I(f) = \frac{1}{8\pi}\frac{\rho}{c} V_0^2 (2\pi f)^4 \quad [10]$$

Although V_0 might depend on frequency, there is normally a physical limit on V_0 resulting from the construction of the device and V_0 tends to a constant as the frequency falls. This means that at low frequencies, the power drops as the fourth power of the frequency, that is, at 12 dB per octave. This is why marine seismic vibrators are said to lack low-frequency output.

In addition to the lack of low-frequency output, small (compared with wavelength) marine seismic vibrators also suffer from a generic problem of impedance matching. Because the source is small compared with the wavelength, the pressure and volume have only a very small phase difference. This presents the drive mechanism with a highly reactive loading. This means that significant power is needed in the device but little of it is radiated as sound. The use of larger devices, or arrays, can reduce this problem.

There are several mechanisms used to drive the volume expansion needed by a marine vibrator. Figure 18 shows a hydraulically driven device where the end of the cylindrical drums moves axially in and out. Marine seismic vibrators offer intriguing possibilities for source development because their waveform is controlled. In the past, they have generally been operated using sinusoidal frequency sweeps, but in principle, any desired waveform could be emitted.

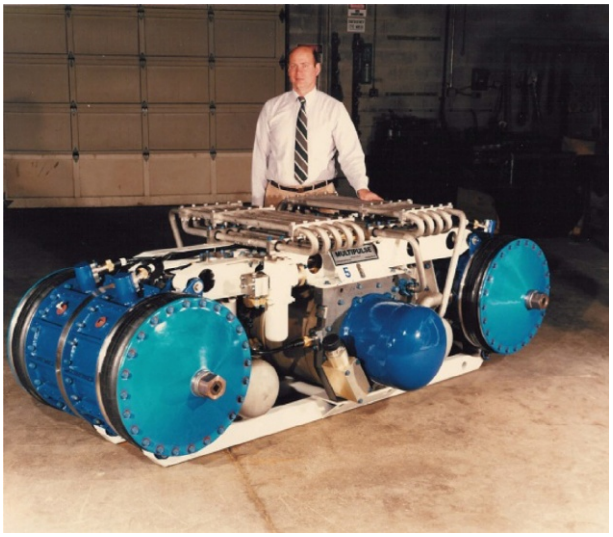


Figure 18 A hydraulically driven marine vibratory source. Photograph provided courtesy of Hydroacoustics Inc.

Figure 19 The seismic vessel on the right is towing two source arrays and eight streamers. From left to right, the ten floats visible on the surface are streamer 1 deflector, streamer 2 deflector, source 1 float 1, source 1 float 2, source 1 float 3, source 2 float 1, source 2 float 2, source 2 float 3, streamer 7 deflector, and streamer 8 deflector. Source 1 has just been fired. Source 2 was fired ten seconds before; the residual bubbles can be seen in the water. Streamers 3–6 are deflected using underwater cables to the other deflectors.

11.06.4.5 Steering, Navigation, and Positioning

Modern seismic vessels are steered by computer so that the sources and receivers follow, as closely as possible, their preplanned positions. Typically, the source will be fired 'on position.' This means that the triggering is determined by the condition that the source is at its closest approach to the desired position. Because the vessel speed is not exactly constant, this means that the time interval between shots is not exactly constant. A less frequently used alternative is to shoot 'on time' where the source is fired exactly at the predetermined time, but the spatial position is allowed to vary slightly.

The primary positioning system used for towed marine surveys is GPS. There is a GPS system on the vessel and one on the tailbuoy at the end of each streamer. In addition, further GPS receivers are usually placed on many of the floats at the head of the receiver array and on the source floats. Figure 19 shows an aerial photograph of a typical 'front end' of a seismic array. The main difficulty is in determining the positions of the streamers, which are below the surface and therefore unable to use GPS. This is done by using a network of acoustic rangers along the whole or part of the length of the receiver array (Bittleston et al., 2000) together with hydrostatic depth sensor and, possibly, magnetic compasses and acoustic depth sensors.

For seabed systems, the positioning problem is simpler because the receivers do not move. A support vessel using

GPS at the surface and acoustic ranging can then be used to locate the receiver nodes.

11.06.5 Illumination

11.06.5.1 Seismic Survey Configurations Geometries

Marine seismic data were originally acquired in the so-called 2D geometry in which a ship tows a source followed by a streamer directly behind it. Depending on the depth to the target, the streamer can be as short as a few kilometers and as long as 6–8 km or more in order to capture as much as possible the reflected energy up to critical reflection angles.

The 3D marine seismic method started to be widely used in the 1980s. Figure 1 illustrates the configuration of such an acquisition system: a ship tows a number of parallel streamers (anywhere between 6 and 20), which can be as long as 6–8 km or more depending on the depth to target. The streamers are towed 50–200 m apart. At the front of the spread, one or two sources are located (in Section 11.06.5.2.1, we will discuss how two sources can be fired in a flip–flop shooting pattern to better populate CMP bins in the inline vs. crossline directions). The 'conventional' 3D marine seismic method allows cost-effective acquisition of data in uniform CMPs and is typically acquired by sailing up and down across a survey area in parallel lines. Since the turning radius of a vessel towing the full receiver array is several kilometers, adjacent lines are typically not acquired consequently but rather in a 'race track' pattern that moves up one sail line at a time.

As discussed in Section 11.06.2.3, purely focusing on CMPs does not guarantee the acquisition of a data set acquired uniformly with respect to azimuth over the target. For a conventional 3D marine seismic acquisition geometry, the azimuth corresponding to the acquisition direction will hugely dominate the acquired data. For a subsurface with anisotropic properties or significant lateral variations, it is therefore highly desirable to acquire data that are rich in both offsets and azimuths over all subsurface locations. As a result, the so-called multi-azimuth (MAZ), rich-azimuth (RAZ), and WAZ acquisition geometries were developed in the 2000s and were demonstrated to provide superior subsurface images over complex targets such as salt structures (see, e.g., Michell et al., 2006). MAZ, RAZ, and WAZ geometries may involve one or several vessels equipped with streamer spreads and sources, shooting into each other's spreads and acquiring data in different directions over the same subsurface locations. Coil shooting (Brice et al., 2012) involves sail lines where one or several vessels acquire data in circles slightly offset to each (coils). The acquired data will be particularly rich in azimuths and offsets over subsurface locations where coils intersect.

11.06.5.2 Seismic Sources

11.06.5.2.1 Single-source and dual-source flip–flop

The crossline separation between CMP points is inherently larger than the inline separation because the lateral separation of the streamers is much larger than the inline separation of the sensors in the streamers. A trade-off between inline and crossline CMP density is achieved by alternating shots to port and starboard ('flip–flop'). This reduces the crossline CMP spacing at the cost of increasing the inline spacing as illustrated in Figure 20. The CMPs each lie, by definition, midway between the source and the receiver. If the streamers are, say, 100 m apart symmetrically disposed behind the vessel, then a single source firing every 25 m would result in an inline CMP interval of 12.5 m and a crossline interval of 50 m. On the other hand, if the source is fired in the flip–flop mode, alternating port and starboard and displaced 25 m from the sail line, then the inline and crossline CMP intervals are both 25 m.

The advent of streamers that sense crossline acceleration (or pressure gradient; Robertsson et al., 2008) may remove the need for this compromise since they provide an improved lateral sampling of shot gathers.

11.06.5.2.2 Simultaneous sources

Seismic image processing was developed on the premise that the basic experiment was to measure the response to a single source that is localized in space and time. Several authors, for example, Berkhout and Verschuur (2012) and Wapenaar et al. (2012), have urged that this approach should be abandoned and the input wavefield, in all its complexity (including simultaneous sources and multiples), should be properly imaged and not approximated by a set of separate point source seismograms. However, at present, the simultaneous source data are separated and then processed as if separate sources had been used. The difference between the 'single shot' data and that obtained by separating a simultaneous source data is known as 'cross talk.' Beasley (2008), Moore et al. (2008), and Hampson

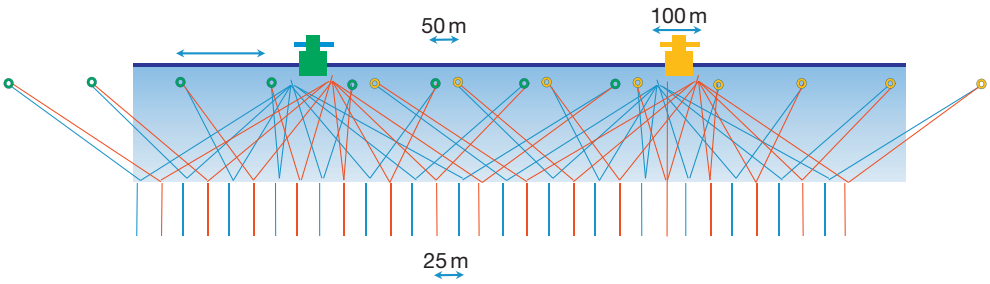


Figure 20 A typical 'flip–flop' configuration. Streamer separation is 100 m; flip–flop source separation is 50 m. Two lines are shown (green and orange), and the line separations are chosen to leave a 25 m offset between the streamer positions. The resulting CMP positions are spaced at 25 m crossline.

et al. (2008) are good general references on the topic. Simultaneous sources have a wider acceptance in land seismic (e.g., Bagaini, 2006).

The separation might, at its most basic, simply involve doing nothing. This has been proposed for land seismic (Huo et al., 2012). The cross talk is not correctly imaged and will, to an extent, be attenuated as a result. In other methods, the source positions are chosen so that the resulting seismograms are very different and can be separated in the shot domain on that basis (see, e.g., Beasley et al., 1998). Another approach is to make use of the coherence from one shot to the next and to do the separation in a domain other than the shot domain. In one variant of this approach, the relative firing times of the simultaneous sources are varied, perhaps randomly, so that they are not exactly simultaneous and can be separated. This is known as 'dithering' (see, e.g., Mahdad et al., 2011; Moore et al., 2008).

11.06.5.2.3 Azimuthally invariant sources

Conventional processing assumes that the source is a point source, localized in space and time. In fact, the source has finite extent in time (the source wavelet) and in geometry leading to directivity (see example in **Figure 21**). The degree to which source directivity affects the seismic image, and how it can be corrected, has been addressed by many researchers. Loveridge et al. (1984) concluded that directivity was an important issue that had been largely ignored until then and that it was particularly important at large offsets and shallow targets. Brummitt (1988) gave an experimental demonstration of its importance by repeating a seismic line using different source arrays. Saunders et al. (2004a,b) concluded that it was a clearly identifiable issue in time-lapse seismic. Hatton and Haddow (1991) investigated how the crossline directivity of a wide array could mimic the effect of migration. There are several acquisition scenarios where the orientation of the source arrays can vary considerably as the data are acquired. An example of this is when a source vessel is used, which tows no streamers. In such cases, the azimuthal directivity can be especially important.

Seismic data are typically processed as if the source wavefield was a monopole (pressure radiated uniformly as a function of emergence angle from the source). To an extent, the actual directivity of the source can be corrected for in the common-receiver domain provided that it does not vary from shot to shot. Krail and Shin (2005) showed how to remove the dipole part (caused by the sea-surface reflection) of the directivity in the tau-p domain. Fokkema et al. (1990) proposed an angular-dependent deconvolution on common-receiver gathers. Roberts and Goulty (1988, 1990) showed how directional deconvolution could be combined with prestack migration and applied in the common offset domain. Ziolkowski and Lopes (1993) showed how to deconvolve the directivity in FK space in the shot domain assuming that the geology is smooth.

One serious difficulty with directional deconvolution in general is that the data are spatially aliased because the spatial shot interval is typically larger than half a wavelength. This is usually true of the inline shot spacing and almost always true for the crossline shot spacing. In this case, we use some layer cake assumptions about the geology to apply the (approximate) source direction deconvolution processing. One overall solution to this problem is to design the source array so that its radiated field resembles that of a monopole, at least limiting the range of takeoff angles that contribute to the image, so that directional deconvolution is not needed. Hopperstad et al. (2001, 2008a,b,c) have shown how to do this by using a mix of deployment depths and spatial distributions of the source elements such that the dipole effect of the sea surface is canceled by the other array effects. Such 'directionless' sources either can be made merely azimuthally independent or can be made to approximate a monopole over all solid angles that contribute to the seismic image.

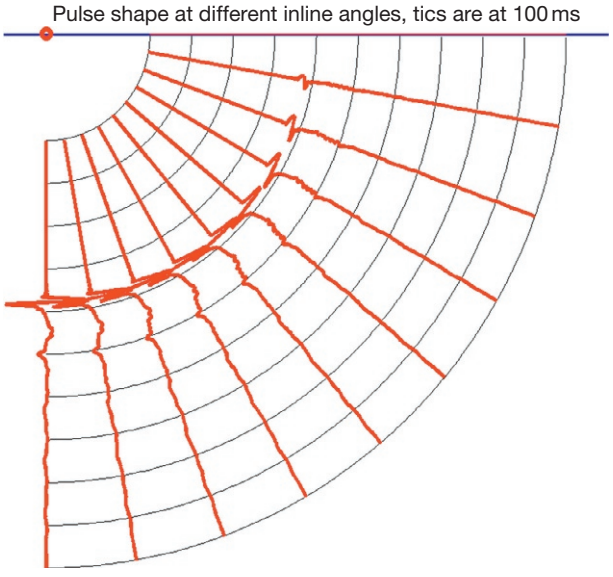


Figure 21 Directivity of an example array showing the far-field signature in various radiation directions.

11.06.6 Sampling and Spatial Aliasing

One of the fundamental limitations in 3D marine seismic acquisition is related to insufficient spatial sampling of the recorded data. Shannon's sampling theorem (also known as the 'sinc interpolation formula') states that a pressure wavefield p at a point y_0 can be exactly reconstructed from its uniformly spaced (infinite number of) samples when the sampling rate σ is greater than the Nyquist sampling rate (twice the spatial bandwidth of the wavefield):

$$p(y_0) = \sum_{m=-\infty}^{\infty} p\left(\frac{m}{\sigma}\right)\mathrm{sinc}(\sigma y_0 - m) \qquad [11]$$

where $\mathrm{sinc}(y) = \sin(\pi y)/(\pi y)$ (Shannon, 1949). Note that for Shannon's sampling theorem to be valid, the Nyquist sampling criterion applies to the shortest wavelength there is in the data, not the shortest wavelength of interest. If the Nyquist sampling criterion is not satisfied and the data are sampled coarser than twice per minimum wavelength, then the so-called aliasing is introduced. Thus, we not only fail to recover the high frequencies in the data, but information folds back (or aliases)

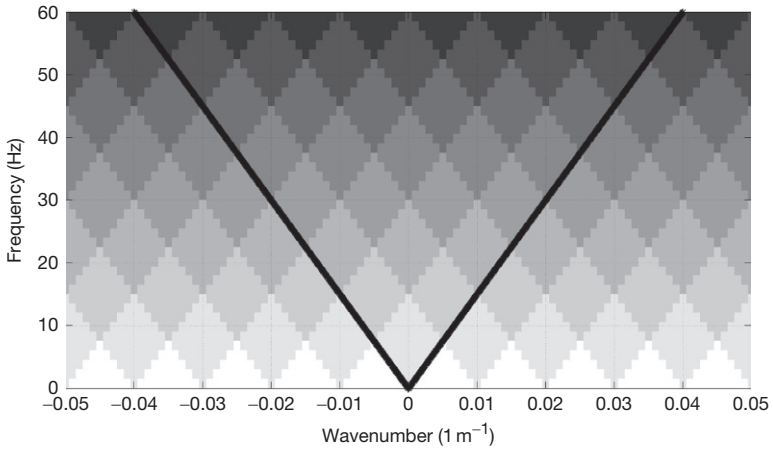


Figure 22 Diagram depicting orders of spatial aliasing in marine seismic acquisition. In this example, the minimum wave-propagation velocity is 1500 m s^{-1}, and the sampling interval is 100 m. Sampling at 100 m generates spectral replicas of the signal cone at regular steps of 0.01 m^{-1}. The replicas overlap in the original signal cone. The intensity of gray shades indicates the order of alias in every region. Reproduced from Vassallo M, Özbek A, Özdemir K, and Eggenberger K (2010) Crossline wavefield reconstruction from multicomponent streamer data: Part 1 – Multichannel interpolation by matching pursuit (MIMAP) using pressure and its crossline gradient. *Geophysics* 75: W853–W858.

into the low frequencies, distorting the baseband spectrum. Aliasing thus limits the ability to reconstruct data in between sampling locations. Figure 22 shows the effect of aliasing for evenly spaced recording locations 100 m apart (strictly speaking, this picture is only valid for an infinite number of sampling points) in frequency wave number space. The solid black lines through zero wave number represent the so-called signal cone. For towed marine seismic data recorded along a horizontal line of receivers, no energy will exist outside the signal cone since no propagating energy can have an apparent velocity lower than that of water (1500 m s^{-1}). Arrivals with this slowest apparent velocity correspond to waves propagating inline with the recording line. In contrast, a vertically incident plane wave (or a horizontally incident plane wave broadside to the receiver array) will appear at all recording locations simultaneously and will therefore have an infinite apparent velocity.

The diamond-shaped patterns in Figure 22 represent different orders of aliasing as different parts of the frequency wave number space are folded on top of other parts. The light gray diamond represents unaliased data, whereas the increasingly darker shades of gray and black show areas of the frequency wave number space where first order, second order, etc., aliasing is present. Signals located in these parts of the frequency wave number space will be contaminated with other signals at other wave numbers that now constitute severe noise. In the example in Figure 22, with 100 m-spaced sampling points, aliasing will occur already at 7.5 Hz.

Ideally, we would therefore like to see marine seismic surveys with a carpet of sources spaced according to the Nyquist sampling criterion for the minimum wavelength of interest (maximum frequency of interest). For each source position, we would also like to see a carpet of receivers with a similar receiver spacing following the Nyquist sampling criterion. In marine seismic surveys, we may be interested in frequencies as high as 150 Hz. Given that the propagation velocity in water is around 1500 m s^{-1}, we will encounter true wavelengths of 10 m, which, for those waves traveling horizontally along the array, requires sensors and sources to be spaced 5 m apart. Clearly, it is both uneconomical and unpractical to conduct such a survey, and various compromises must be made in the acquisition geometry, in particular with respect to how dense sources and receivers are sampled. In practice, the situation is not quite as bad as what we just outlined. Depending on the situation, we may obtain adequate results even in the presence of some aliasing. A more pragmatic sampling criterion should therefore be determined by the spatial sampling required for a certain multichannel process (Abma and Kabir, 2005). The specific multichannel process may be as simple as stacking to reduce coherent noise or as complex as prestack depth migration. However, even with this less strict sampling criterion, 3D marine seismic data are often plagued with aliasing on both the source and the receiver sides that degrades the quality of the final seismic product.

For towed marine acquisition configurations, several streamers are towed behind a vessel. Sensors are normally spaced within the Nyquist sampling criterion along the streamer. However, the distance between streamers is typically much greater. Although surveys have been carried out with streamers as densely spaced as 37.5 m apart (Long, 2004), 75–100 m streamer separations are much more typical. Long (2004) noted that "a significant component of the seismic noise contaminating 3D images actually arises during processing, as an unfortunate, and inescapable artifact from poor 3D spatial sampling. When the crossline acquisition dimension is sampled at an equally small interval as the inline dimension, a much larger frequency bandwidth than typical of standard 3D acquisition can be preserved throughout all stages of processing, free of aliasing, and free of artifacts."

For ocean-bottom cables, the situation is somewhat worse. Existing systems have an inline spacing that is typically 50 m or at best 25 m. Crossline spacing can be several hundred meters. The situation is even worse for ocean-bottom nodes, which are typically spaced several hundred meters apart, in both the inline and crossline directions.

Source-side sampling is often similar for towed marine, ocean-bottom cable and node surveys. As discussed previously,

the inline sampling is governed by the record length to capture all events of interest while minimizing residual shot noise and also by the speed of the vessel towing the source. The inline source spacing can be as small as 25 m but is more typically 37.5 or 50 m. Two source arrays are often deployed from a single-source vessel, spaced some 50 m apart laterally. As described earlier, 'flip–flop shooting' results in a more uniform and denser sampling of CMPs at the expense of sparser inline source spacing.

In a seabed seismic survey, it is often of interest to record shear wave data. Because of the lower propagation velocity of shear waves, the record length and, correspondingly, the source-firing interval are longer. In addition, adjacent source lines may be spaced 50 m to several hundred meters apart. Ocean-bottom cable geometries are most typically carried out in a crossline geometry similar to many land seismic surveys where source lines are perpendicular to receiver lines, resulting in more evenly spaced data in the CMP domain.

Aliasing or gaps in the data may also be introduced by deviations from the desired acquisition geometry caused by, for instance, currents resulting in streamers that deviate from the desired positions (the so-called feathering). In such cases, it is often required to reshoot specific parts of the survey area for 'infill.' Furthermore, it is not unusual to find surface obstructions, such as oil platforms, so that it is not possible to acquire seismic data uniformly throughout the entire survey area. Ocean-bottom cables or seabed nodes offer practical solutions to this problem that sometimes make them an attractive alternative to towed marine seismic acquisition.

11.06.6.1 Receiver-Side Techniques

The simplest technique to reduce aliasing is the so-called normal-moveout correction. This technique has been used since the beginning of reflection seismic processing to construct images of the subsurface. The basic idea is simple: by assuming that the data consist of primary reflections from subsurface interfaces in a horizontally layered Earth, events at a specific offset trace are shifted to the time in the trace where they would have been if the data trace would have been acquired at zero offset. Once this moveout correction has been applied for a receiver line and a shot location, the data traces can be summed together to produce a single stacked trace with the enhanced signal-to-noise ratio (Yilmaz, 2001). Note that the NMO process is a nonlinear process as it stretches the trace causing a frequency shift and is by no means an ideal dealiasing method. Nevertheless, the approach can be used for dealiasing by interpolating data after NMO followed by an inverse NMO to remove most of the undesired nonlinear effects.

However, if Earth properties varied only with depth and a 1D approximation always worked, we would not need 3D seismic data in the first place. In reality, the most interesting seismic imaging problems from an exploration point of view involve structures that do not satisfy these criteria. As a result, a wide range of sophisticated interpolation techniques have been developed to dealias seismic data. Abma and Kabir (2005) compared the pros and cons of a range of popular techniques for interpolation of seismic data. In order to overcome aliasing, the techniques make certain assumptions on the character of the seismic data. Several of the most popular techniques assume linear events, in particular the technique by Spitz (1991). In Spitz's method, a gather of traces in the domain where they are to be interpolated is Fourier-transformed in time. Starting from the unaliased low frequencies, higher-frequency (aliased) components are reconstructed by assuming a projection of the lower-frequency components following the assumption of linear events. In the most general case, the surface seismic interpolation problem is referred to as being five-dimensional: two receiver location coordinates, two source location coordinates, and the time–frequency axis. High-order interpolation techniques have therefore been of particular interest to researchers in recent years as computational resources have enabled the implementation of sophisticated but computationally intensive algorithms (Trad, 2009).

A radically different solution to the aliasing problem was introduced by Robertsson et al. (2008) who realized that recordings of the 3D particle velocity vector in addition to the pressure relax the Nyquist sampling criterion by a factor of two or more. The basic principle follows from the equation of motion that states that the spatial gradient of pressure data, ∇p, is proportional to the acceleration of particle motion, $\ddot{\boldsymbol{u}}$:

$$\nabla p = \rho \boldsymbol{u} \qquad [12]$$

where ρ is the density of water. Intuitively, the spatial gradient of the pressure wavefield will be useful for understanding the structure of the wavefield away from the recording location. Robertsson et al. (2008) showed that recording the horizontal components of particle acceleration in addition to the pressure results in a special form of Linden's multichannel sampling theorem (Linden, 1959):

$$p(y_0) = \sum_{m=-\infty}^{\infty} \left\{ p\left(\frac{2m}{\sigma}\right) + \rho\left(y_0 - \frac{2m}{\sigma}\right) \ddot{u}_y\left(\frac{2m}{\sigma}\right) \right\} \mathrm{sinc}^2\left(\frac{\sigma y_0}{2} - m\right) \qquad [13]$$

where $\ddot{u}_y$ is the component of particle acceleration in the direction of interpolation (proportional to the spatial gradient wavefield in the crossline direction as shown by eqn [12]). Equation [13] is a 'sinc interpolation formula' for a multicomponent streamer (recording a minimum of pressure and crossline components of particle velocity, the latter being simply obtained by integrating the acceleration). It is useful to compare this against the 'sinc interpolation formula' for pressure data only in eqn [11]. In particular, note the factor ½ that occurs with every factor σ in eqn [13] compared to eqn [11]. This difference in the two sinc interpolation formulas is fundamental and reflects the fact that the Nyquist sampling criterion has increased by a factor of 2 in eqn [13]. Vassallo et al. (2010) presented an interpolation algorithm for multicomponent data (pressure and particle acceleration recordings) based on a basis function expansion (a so-called matching pursuit algorithm) that in combination with a linearity assumption, similar to what was proposed by Spitz (1991), provides a powerful engine for dealiasing multicomponent streamer data. Özbek et al. (2010) took this concept even further by additionally letting the ghost constrain the interpolation. Together, all these advances provide a powerful framework for crossline reconstruction of towed marine multicomponent seismic data. Based on the work by Robertsson et al. (2008),

Amundsen et al. (2010) explored the benefits of interpolation of seabed node data where higher-order gradients of pressure data can also be measured, further constraining the interpolation problem and allowing for sparse sampling of the recorded wavefield.

For processing of seabed seismic data and in particular for seabed node acquisition, it is common to use reciprocity arguments for applications such as deghosting, multiple removal, and imaging. Instead of processing common-shot gathers (which may be irregular or too sparse), common-receiver gathers are formed. Although reciprocity can be invoked in such gathers between conventional marine seismic sources (of monopole character) and hydrophone recordings, reciprocity cannot generally be used for other recorded quantities such as particle motion recordings. The literature is full of examples where seabed seismic data is processed by combining pressure and particle motion recordings in common-receiver gathers arguing that this follows from reciprocity. In general, this is not correct unless one assumes a medium that varies with depth only where Green's functions become functions only of the lateral source–receiver offset.

11.06.6.2 Source-Side Techniques

In current towed marine seismic surveying, the spatial shot interval is too large to avoid spatial aliasing. This aliasing (in common-receiver domain) is a result of residual shot noise. In principle, in a conventional towed marine seismic experiment, each shot could be fired immediately after the two-way time to the target from the previous shot has elapsed because the portion of the seismogram beyond the target reflections does not contribute to the image (unless the multiples are imaged, Berkhout and Verschuur, 2012).

With conventional tow speeds, target depths, and frequency ranges, such a shot time interval would, as it happens, provide an alias-free spatial sampling (inline). However, in practice, the source is fired only after enough time has elapsed so that the reverberation from the previous shot (which is not correctly imaged) has decayed away to a sufficiently low level. This is illustrated in Figure 23. This time interval (usually 10 s) is typically considerably longer than the target time, and the resulting spatial shot interval (usually 25 m) is too large to avoid aliasing at the higher frequencies. The residual shot noise has been studied by Landrø (2008) who looked at the effect on the seismic image and by Herber et al. (1997) who compared the reverberation to ambient noise levels.

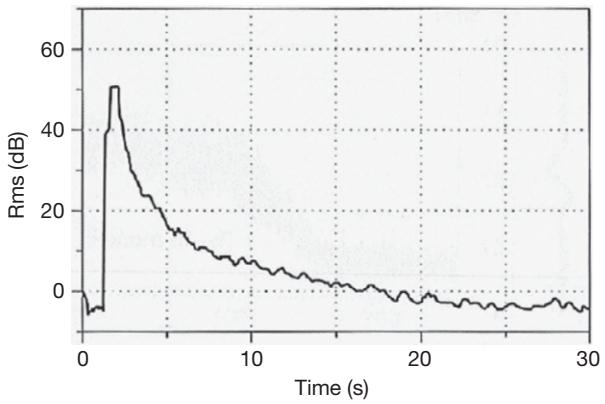


Figure 23 From Landrø (2008), showing the typical energy decay of a seismic shot over 30 s. The time between shots must be large enough that the reverberation has decayed to a sufficiently low level.

One solution to this would be to use a slower tow speed for the survey but that immediately increases the duration and hence the cost of the survey. In addition, systems that have been designed to be towed at one speed cannot readily be adapted to be towed at a much slower speed.

Recently, Robertsson et al. (2006, 2012) introduced methods for acquiring multicomponent data also on the source side for towed marine applications. This requires the generation of closely spaced source points so that the response of source-side gradients (dipole-type sources) can be generated. The concept is a topic of ongoing research but is likely to be the only way forward if one wants to genuinely process seismic data solving wavefield separation and multiple or imaging problems on the source side.

11.06.7 Bandwidth

The temporal bandwidth of the received seismic signal is the range of useable frequencies, from a minimum frequency to a maximum frequency, measured in units of hertz. It defines the vertical resolution in the image, for example, the ability to separate thin beds. The resolution limit is typically defined to be about a quarter of the dominant wavelength of the signal (two-way propagation over the so-called Rayleigh resolution criterion), and vertical resolution is improved by including higher frequencies and broader bandwidth. Figure 24 shows the relative contributions of high and low frequencies to the seismic wavelet. The left panel has increasingly more high frequencies from top to bottom, showing how these reduce the central lobe width and hence improve the resolution. The right panel has increasingly low frequencies from top to bottom, showing that while the central wavelet lobe width remains the same (governed by the high-frequency limit), the wavelet-side lobes are compressed, increasing the interpretable resolution of the wavelet.

Seismic data are by nature band-limited because the seismic source only emits a given range of frequencies and anelastic attenuation, which is more severe on the higher frequencies, limits the high frequencies we can receive back from the Earth. Marine seismic methods are additionally band-limited because of the free-surface ghost that is inherent in the method. This is the (almost total) downward reflection of the upward traveling wavefield at the sea surface, which is reversed in polarity and travels back down to the receivers where it is recorded as a delayed version of the upgoing wavefield. The resulting pressure wavelet is extended and contains both the upward and downward traveling components. This causes constructive and destructive interference of the signals at different frequencies depending on the depth of the receivers. Figure 6 shows a schematic of the receiver ghost geometry, the resulting wavelet, and the theoretical frequency spectra of the free-surface ghost for two different recording depths of 6 and 20 m. For a hydrophone recording device, notches in the spectrum, at vertical incidence, occur at $Fn = nv/2d (n = 0, 1, 2, 3, \ldots)$, where v is the speed of sound in the water layer and d is the depth of the streamer below the sea surface.

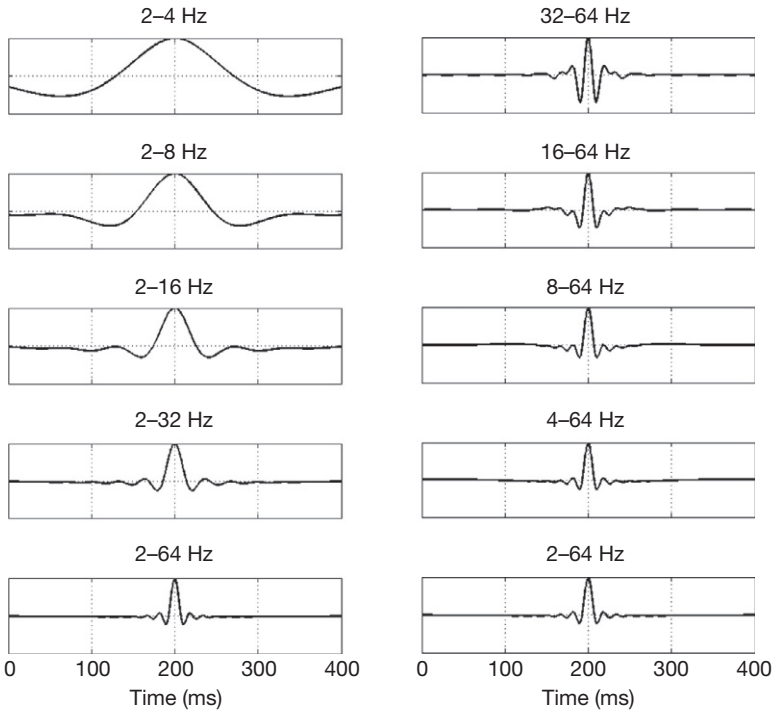


Figure 24 The relative contribution of high frequencies (left) and low frequencies (right) to the seismic wavelet resolution.

The ghost notch frequency increases with incidence angle (offset) because the path length difference of the direct and ghost event decreases. For a given angle, $\emptyset$, $Fn(\emptyset) = nv/2d\cos(\emptyset)$. A further free-surface ghost is also generated at the source, which is also towed at depth of a few meters below the sea surface, further compounding the limitation on bandwidth.

Maximizing the useable bandwidth has always been one of the goals of data acquisition and processing, and this must be achieved while maintaining an acceptable signal-to-noise ratio in the data. What is acceptable is governed by the intended application of the survey. For example, a time-lapse monitoring survey will require a greater signal-to-noise ratio to see small changes in the reservoir than will a structural image for exploration purposes.

In recent years, extending the bandwidth to include more low frequencies has become increasingly important. While it has always been understood that low frequencies are required for good penetration and imaging at depth, the low frequencies are also now realized to be important for wavelet stability in inversion methods such as acoustic impedance inversion and for accurate forward models within FWI methods: for FWI to work, the starting model must yield seismic data, which are kinematically within half a period of the observed data; otherwise, the velocity updating goes in the wrong direction and the solution gets trapped in a local minimum. By including and starting the inversion with low frequencies, a FWI algorithm can be designed so that the local minima are avoided.

The following sections describe a number of acquisition methods designed to increase the data bandwidth, many of which have only become accepted commercially within the last few years at the time of writing.

11.06.7.1 Receiver-Side Techniques

The seismic bandwidth is largely determined by the receiver system, and careful design of this component of the acquisition system can make a significant difference to the useable bandwidth.

From our discussion of the free-surface ghost, we can see from Figure 6 that towing streamers at a shallow depth favor the higher frequencies at the expense of attenuating the low frequencies, whereas towing streamers at a deeper depth favor the lower frequencies, at the expense of attenuating frequencies within the desired seismic bandwidth. How to overcome this ghost effect has been the subject of geophysical research for many years. Over the last ten years or so, a number of commercial 'deghosting' solutions have appeared on the market.

11.06.7.1.1 Single-streamer deghosting methods

We include a short introduction here on 'single-streamer deghosting.' What we mean are methods that are applied to conventional (constant depth) towed streamer acquisition. This section is included for completeness as it should be understood that single-streamer methods shape the seismic wavelet (for stability, resolution, and interpretability) but do not improve the signal-to-noise ratio. In particular, they do not fill the ghost notches as there are no additional data. The methods are akin to statistical or deterministic deconvolution processing methods that attempt to broaden the data bandwidth (see, e.g., Yilmaz, 2001). Single-streamer deghosting methods are deterministic methods and use knowledge of the streamer depth to estimate the ghost operator and then design an inverse operator to correct for the ghost effect. The skill is then in optimally band limiting the data to control the noise, which is inevitably amplified at the low and

high frequencies around the ghost notches. The deghosting operators can be designed to account for the roughness of the sea surface, which essentially modifies the height of the water column (and hence ghost notch frequency) along the streamer (Amundsen et al., 2005; Robertsson and Kragh, 2002). To do this, a model of the sea surface must be known for all times during the seismic record. Laws and Kragh (2002) described the effect of the rough sea on the seismic wavelet, and Laws and Kragh (2006) showed a method to derive the necessary sea-surface shape. Stabilizing the wavelet for the perturbations (errors) caused by the rough sea ghost can be important in time-lapse seismic for reservoir monitoring (Laws and Kragh, 2002). Kragh and Laws (2006) described why this is difficult to achieve with conventional processing methods.

11.06.7.1.2 Over–under methods

The over–under deghosting method (e.g., Hill et al., 2006; Sonneland et al., 1986) deploys streamers at two different cable depths and at the (ideally) same lateral position, that is, as vertically aligned streamer pairs. Choosing the cable depths such that the ghost notches are complimentary to each other (i.e., where one streamer has a notch, the other has a peak) allows the notches to be filled by essentially using data from either cable where there is no notch. While the notch at zero frequency still exists, towing both cables relatively deep significantly increases the recorded amplitude at the low frequencies due to the steeper ghost notch roll-off (e.g., **Figure 6**, blue 20 m depth spectra in the right-hand panel). In addition, towing the cables deeper means they are further away from the sea surface and the noise which the wave motion can impart on the streamers (see **Figure 5**). We should note here that lowering the noise floor is equivalent to increasing the signal strength.

Although the over–under method was published in the 1980s (e.g., Sonneland et al., 1986), the method was not widely accepted due to the difficulty of controlling the streamers at the different depths due to crosscurrents. This was overcome two decades later with the advent of lateral streamer steering, and Hill et al. (2006) and Moldoveanu et al. (2007) gave examples of successful application to bandwidth extension. However, while the over–under method is able to provide high-quality broadband data, the reduction in the 3D spread width due to the need for two streamers at each lateral position meant that the method was half as productive (roughly translating to twice the acquisition cost) as a traditional single depth spread. A further development of the method (Kragh et al., 2010) considerably relaxed this constraint by only using a smaller number of 'under' streamers, which were used for recording of the low frequencies only, while the upper spread recorded the full-bandwidth data. Interestingly, the over–under method had been considered a long time previously, for example, an early reference is found in a patent filed by Haggerty (1954).

11.06.7.1.3 Variable tow depth methods

Variable tow depth methods acquire data using a slanted cable or variable depth shape along the cable. The basic idea is that each offset has a different depth and therefore has a different ghost notch frequency (introducing the so-called notch diversity). After processing the data, where essentially all offsets are summed at the imaging step, the ghost notches are filled providing a broadband solution. The cable shape can be tuned to optimize the notch diversity for the resulting final bandwidth, though, as in the over–under solution, the notch at zero frequency will remain. This is again improved by the deep-tow aspect of this acquisition method. Clearly, this method reduces the overall signal-to-noise ratio at each frequency as there are only a limited range of offsets contributing to any given output; however, the data recorded to date suggest that this is not limiting. The original idea dates back many years (Ray and Moore, 1982) but was reintroduced more recently (Soubaras and Dowle, 2010) including data-specific imaging methods.

11.06.7.1.4 Multicomponent towed streamer methods

The discussion on ghost responses so far has been limited to hydrophone receiver systems. Hydrophones have been used in towed marine seismic cables almost exclusively until 2007 when the GeoStreamer® system (Long et al., 2008) was introduced. This two-component system contains a vertical velocity measurement obtained using gimbaled geophones mounted in the cable. The theoretical ghost notches for velocity are different to those for pressure (hydrophones) because the velocity sensor senses the direction of motion of the wave, and therefore, even though the polarity of the downgoing wavefield is reversed, its direction is also reversed. This results in the ghost reflection being recorded with the same *apparent* polarity as the upgoing wavefield. Theoretical (vertical incidence) notches now appear at $f_n = nv/4d$ (where now $n = 1, 3, 5, \ldots$). **Figure 25** shows spectra for both pressure (hydrophone) and vertical velocity recordings at 20 m depth. Note how the peaks and troughs (notches) are complementary. This forms the basic concept of multicomponent deghosting methods that have been used for seabed seismic data (see the succeeding text) for decades. The data are essentially summed resulting in a spectrum with no notches (e.g., Tenghamn and Dhelie, 2009). Note that with this multicomponent deghosting solution, the notch at 0 Hz (DC) is also removed in theory. This requires high-fidelity velocity

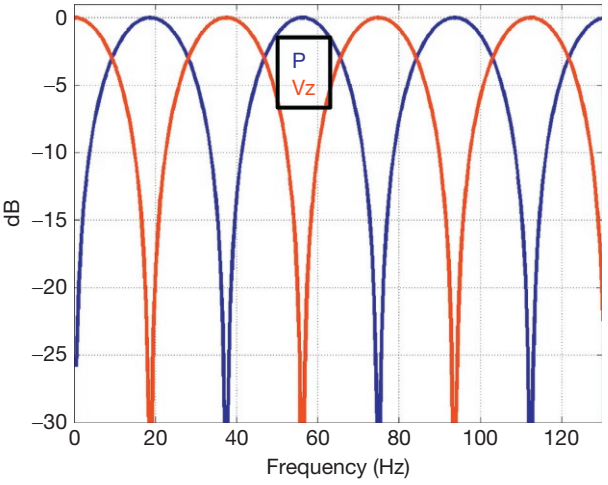


Figure 25 Resulting spectra for vertical incidence ghost reflections for pressure (hydrophones) in blue and vertical velocity in red, for a depth of 20 m. Note the complementary peaks and troughs from the two different measurements.

measurements right down through the low-frequency bandwidth to be acquired (Caprioli et al., 2012).

11.06.7.1.5 Seabed seismic techniques

Placing the receivers on the seabed or seafloor will also provide broadband seismic data. However, this acquisition method is distinctly different from towed streamer methods as seabed systems will record the full elastic wavefield, more akin to land seismic data. The seabed receiver systems can be either cables or individual nodes, but in both cases, they usually record three-component motion elastic data (the wavefield in the seafloor) plus the acoustic pressure in the water just above the seafloor, using an additional hydrophone. It is the combination of the hydrophone and geophone data that allows for the separation of the upgoing and downgoing wavefields (Barr and Sanders, 1989). This removes the water-layer reverberations, which in the case of seabed data include the ghost reflection, providing the broadband solution. With the knowledge of the seabed material properties, the elastic wavefield can be further decomposed into P- and S-wave components (see, e.g., Wapenaar et al., 1990).

11.06.7.2 Source-Side Techniques

Historically, the design of marine seismic sources concentrated on producing a vertically traveling far-field signature (pulse shape) that had a broad, flat spectrum. The metric often used was the primary-to-bubble ratio. This is the ratio of the peak–peak amplitude of the initial part of the pulse to the amplitude of the later 'bubble' oscillations. It is a proxy for spectral flatness. Parkes and Hatton (1986) gave more details of such metrics.

More recently, this requirement for a flat spectrum has been challenged on the basis that it does not result in a seismic image that has uniform signal-to-noise ratio at all frequencies. Laws et al. (2008) and Kragh et al. (2012) studied the SNR of the seismic image by acquiring and processing to an image state repeated seismic lines, including lines where no air guns were fired (only noise, no signal). They concluded that the resulting SNR is typically wanting at low frequencies and often excessive at higher frequencies. It is worth noting that the 'noise floor' in the seismic image does not result from ocean ambient noise; it originates from the self-noise of the acquisition system, which is itself affected by the weather. If streamers could be made quieter, then seismic sources could become quieter.

Air gun array spectra are designed to have a smooth appearance by using a mix of air gun volumes (and thereby a mix of bubble oscillation frequencies). In addition, clusters of air guns are often used within the array. A cluster has a period corresponding to the total volume of air, but it is more heavily damped. The total volume of air available is typically fixed for a particular vessel, and the question is how best to distribute this among the bubbles to give the best spectrum (Giles and Johnston, 1973). In the past, seismic companies would select the mix of volumes based on experience and experimentation. With the advance of modeling codes (Laws et al., 1990; Ziolkowski, 1970; Ziolkowski and Metselaar, 1984), a more rigorous process of designing air gun arrays is now adopted (e.g., commercial modeling packages such as 'Modgun' and 'Gundalf').

Hegna and Parkes (2011) showed that at low frequencies (below 10 Hz), a typical array spectrum does not depend on the array depth. The reason is that the increase in output caused by the reduced cancelation with the sea-surface ghost is almost exactly canceled by the reduced output caused by the change in bubble oscillation period due to the increased hydrostatic pressure. Hopperstad et al. (2012) went further and showed that the far-field spectrum may be usefully considered in two parts: above and below the lowest bubble resonance frequency in the array. Below this frequency, the output spectrum depends only on the total volume of air released (and its pressure); it does not depend on the way in which the air is divided up among the guns or how the air guns are distributed with depth. This is illustrated in **Figure 26** where the same fixed volume of air is released in a wide variety of array configurations. Below 10 Hz, the output is the same.

The high-frequency end of the spectrum is typically increased by using more air guns. Vaage et al. (1983) showed that the peak amplitude of an air gun pulse is roughly proportional to the firing pressure and independent of the volume; the height (not width) of the peak is dominated by the high frequencies. Since the peak pressure of the array is the sum of the peak pressures of the individual bubbles, it follows that the peak of the array is proportional to the number of guns, for a fixed total volume of air.

In order to increase the low-frequency output of the array, there are only two options: increase the total quantity of air and lower the minimum bubble oscillation frequency. Hopperstad et al. (2012) proposed doing the latter by grouping the air guns sufficiently close that the coupling caused them to oscillate at a lower frequency, but not so close that the clustered damping effects (see earlier) become important. Such 'hyperclusters' have been successful in increasing the low-frequency output of the array without using more air. Another attempt to increase the low-frequency output is termed 'bubble tuning' and is described by Lunnon et al. (2003) and Christie et al. (2006). In this case, the air guns are timed so that instead of the initial primary peaks being aligned in time, it is the first peak of the bubble oscillations that are aligned.

Another consideration in source bandwidth is the effect of the sea-surface ghost. In the past, it was usual to deploy all the air guns at the same depth and to choose that depth so that the ghost notch frequency (at which the downgoing reflected wave cancels the direct) was just above the band of interest (7.5 m was a typical depth). However, it is now becoming more usual to deploy arrays with several depths. By doing this, broader spectra can be obtained; see, for example, Cambois et al. (2009), Moldoveanu (2000), and Hegna and Parkes (2011). With marine vibrators, there can be the additional advantage that units could be placed at a depth most appropriate to the emitted frequency (ideally ¼ wavelength) and could even be moved vertically as the sweep progresses (Laws and Morice, 2007).

11.06.8 Time-Lapse and Reservoir Monitoring

Time-lapse seismic is the term used to describe the process of repeating a seismic survey over the same geographical location but at a later time. Differences in the resulting images between

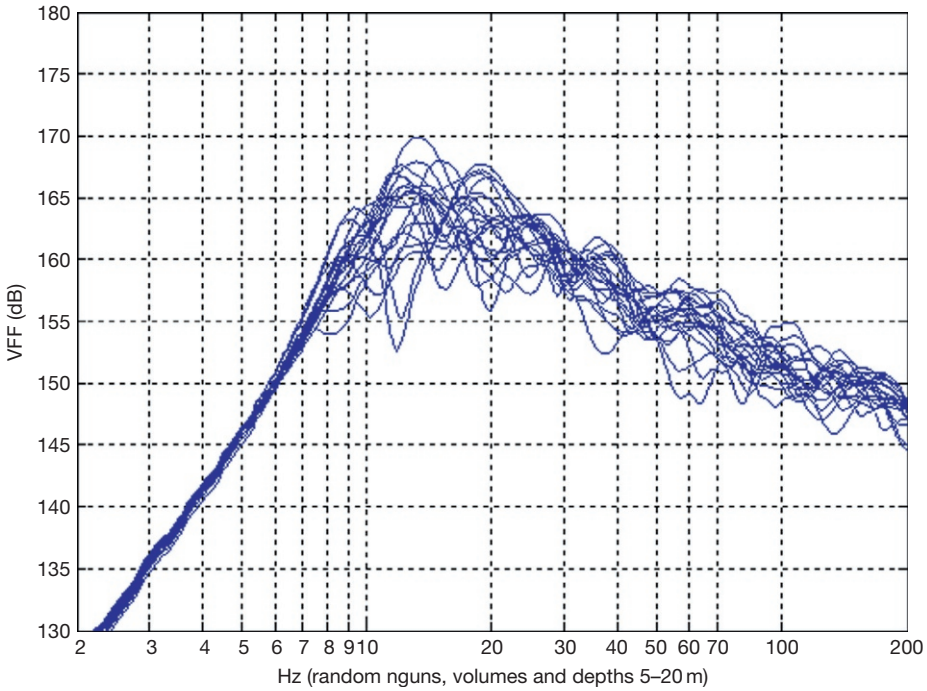


Figure 26 Random arrays with the same total volume showing that the low-frequency output (below the lowest bubble resonance frequency) is unaffected by the details of array design.

the base survey and the monitor survey can be attributed to changes in the reservoir over time, and of particular interest are oilfield production-related changes such as fluid movement. If the individual surveys are 3D surveys (which is typical), then the term '4D' seismic is often used to describe this process, time being the fourth dimension. Multiple-repeat surveys are now common place to monitor the progress of reservoirs as they mature. This is reservoir monitoring (e.g., Fayemendy et al., 2012). Monitoring the progress of a reservoir during production can have huge impacts on the management and production process. Compartmentalization and bypassed oil can potentially be identified, production and drilling well placement plans can be modified, and production rates can potentially be significantly accelerated.

The first seismic time-lapse experiments were laboratory-based, and early time-lapse field tests were carried out on steam injection projects in the late 1980s and early 1990s. Greaves and Fulp (1987) documented one of the earliest field examples and Lumley (2001) gave a nice introduction and history as did Jack (2001). These early time-lapse experiments were land-based with most of the land examples being typically shallow reservoirs/heavy oil and including some novel monitoring techniques (e.g., Forgues et al., 2006; Meunier et al., 2001); however, the majority of current large-scale commercial time-lapse seismic surveys are marine-based.

It should be noted that while 3D repeat surveys are typical on the high-end level of the seismic market, all forms of seismic surveying can be repeated and lend themselves to time-lapse studies, for example, seabed surveys (Ceragioli et al., 2010), downhole VSP surveys (Campbell et al., 2011; Landrø, 1999), and crosswell seismic surveys (Spetzler et al., 2008).

To achieve a successful time-lapse seismic result, the differences in the surveys caused by effects that do not reflect the subsurface changes that are being monitored must be minimized. This 'residual time-lapse noise' depends on many factors such as source and receiver positions, source signature variations, and coupling conditions and on the fidelity or accuracy of the recording equipment itself, which is typically not repeatable. Kragh and Christie (2002) reviewed non-repeatability and looked in detail at the two metrics, the normalized root mean square (NRMS) difference and the predictability, which are now routinely used to measure the time-lapse repeatability. Eiken (2005) gave a nice review of factors influencing the seismic repeatability. Repeatability metrics such as NRMS are highly sensitive to the smallest differences in the seismic measurements. Laws and Kragh (2002) showed how even the effect of a rough sea surface in marine seismic surveys can influence the time-lapse noise, generating a time-lapse noise of up to 10% or so. However, it is now well established that well-repeated surveys with careful processing can show residuals well below 20% NRMS, for example, Day et al. (2010), Li et al. (2004), and Eiken et al. (2003).

A further consideration is the overburden heterogeneity that plays an important role in defining the accuracy of repeating of source and receiver positions in time-lapse surveys (Calvert, 2005b; Landrø, 1999). It follows from this that positioning repeatability requirements will differ with survey location.

However, high survey repeatability still remains the goal of modern reservoir monitoring (Fayemendy et al., 2012), and **Figure 27** shows a simple demonstration of this, taken from Kragh and Christie (2002). The top panel shows a difference

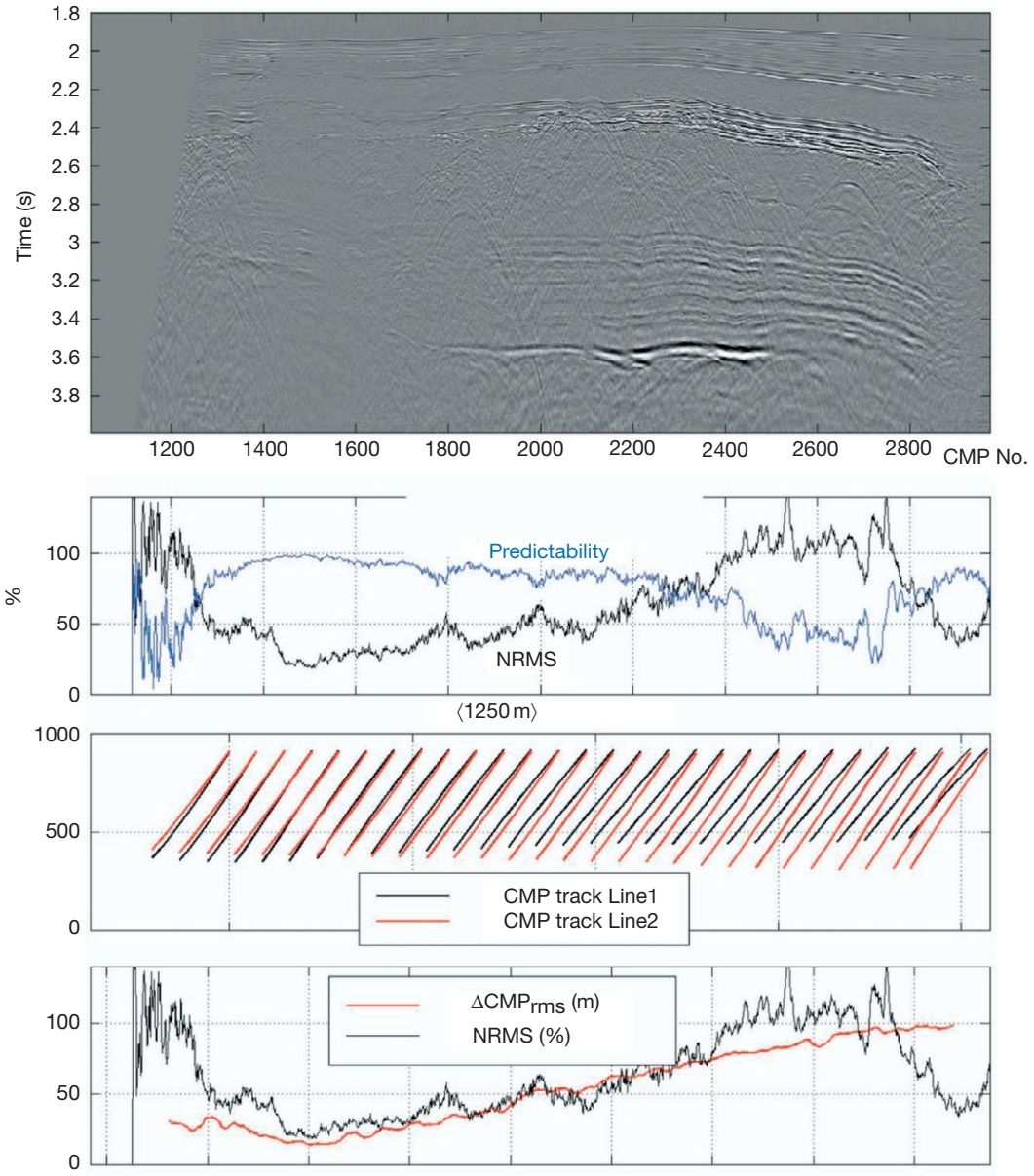


Figure 27 Simple demonstration of the correlation of seismic repeatability and positioning error. Top down: Difference image; NRMS and predictability; CMP tracks for every 30th shot position along the line (black, first pass; red, second pass); NRMS and mean midpoint error. Reproduced from Kragh E and Christie P (2002) Seismic repeatability, normalized rms and predictability. *The Leading Edge* 21: 640–647.

image obtained after subtracting two processed stacked seismic lines. There are clear features across much of the difference image that relate to subsurface structure, but between CMPs 1400 and 1700, the difference is much quieter; the seismic is, visually, highly repeatable. The second panel down shows the repeatability metrics of NRMS and predictability values computed for each CMP over the time window 2–4 s. There is a clear negative correlation between these two metrics, and note that where the data are visually repeatable, the NRMS values are low and the predictability values are high. The third panel down shows CMP tracks (essentially representing the streamer positions) for every 30th shot position along the line with black lines showing the first pass and the red lines showing the second pass. Clearly, the data are most repeatable where the streamer positions most closely coincide. The lowest panel plots the NRMS values (as in the second panel down) with the rms error of the midpoint locations (in meters) in red. A roughly linear correlation is evident, with the most repeatable data corresponding to rms errors in midpoint location of 15–20 m, which, in this case, is mostly a crossline error. This example demonstrates the strong correlation of NRMS with the positioning error.

11.06.9 Conclusions

Marine seismic data are acquired to image the subsurface for a wide range of applications, ranging from shallow engineering

or site surveys (upper 10s–100s of meters) to global scales to understand the structure of the deep Earth and the processes that form it. In this chapter, we focused on describing marine reflection seismic applications related to hydrocarbon exploration and production since these methods are the most advanced and also leading the development of marine seismic methods for other application areas.

Although 2D marine seismic methods are still used, particularly to obtain an initial idea of subsurface structures and complexity in areas where seismic data have not yet been collected, 3D seismic methods completely dominate marine seismic acquisition and processing today. Conventional 3D marine seismic data are typically acquired from a ship towing one or two sources and between 6 and 20 streamers containing sensors recording the seismic reflection data generated by the sources. The streamers are towed some 50–150 m apart and are typically somewhere between 6 and 12 km long. The streamers typically contain a large number of hydrophones recording the pressure fluctuations due to propagating waves, sometimes as densely spaced as 3.125 m. The ship towing the sources and the streamers will acquire adjacent lines of data covering the entire survey area of interest. This will typically result in the acquisition of reflected energy corresponding to a wide range of target incidence angles. However, such data will only contain seismic reflections with propagation paths in the subsurface close to a single azimuth. In areas with strong anisotropy and/or a complex geology (e.g., salt structures), it is often necessary to acquire data that are also rich in azimuths to properly illuminate the subsurface in order to obtain a good image. Such WAZ acquisition techniques have become increasingly common over the last 5–10 years (see, e.g., Michell et al., 2006).

Apart from towed marine acquisition systems, marine seismic data are sometimes recorded with sensors on the seabed deployed as either nodes, retrievable cables, or permanent monitoring installations. Such acquisition systems will typically record pressure as well as three-component particle acceleration or velocity data. The pressure component and the vertical component of particle velocity are used to remove water-layer multiples and to increase bandwidth, whereas the horizontal component of particle velocity predominantly records shear waves converted and reflected from P-waves incident on subsurface reflectors. Although shear waves, in principle, provide valuable complementary information about the subsurface and properties, the recordings are rarely utilized to their full potential. Instead, seabed seismic data are mostly used because they provide broadband (temporal bandwidth) and WAZ data that can be acquired cost-effectively, particularly in areas with surface obstructions.

Marine seismic sources most commonly consist of arrays of air gun elements. In an air gun, air is compressed and quickly released, creating a sudden increase in hydrostatic pressure that results in an acoustic pulse and an oscillating air bubble. The air gun is a remarkably simple and robust device that contains only two moving parts (a solenoid valve and a piston). This is required to ensure robust operations where a source array must be able to fire thousands of times without operational failure. The operational robustness in combination with its adequate efficiency in generating low frequencies is the main reason why air gun sources have remained the marine seismic source of choice for half a century.

Marine seismic data are acquired not only to obtain images of the subsurface but also to image or monitor changes in the subsurface, for example, changes in a hydrocarbon reservoir or its overburden due to production. Such surveys are referred to as time lapse or 4D and place particular demands on the acquisition and processing workflow. In particular, the monitoring survey should repeat the baseline survey as close as possible, in terms of both acquisition parameters and source and receiver locations. The better the illuminating signal (and noise) can be repeated, the better the difference in signals can be detected.

Marine seismic methods are continuously evolving, and the last decade has seen major shifts in technology and its use. Firstly, the need for broader temporal bandwidth has led to the development of acquisition systems and techniques that can better discriminate between signal and noise. In particular, single-sensor recordings allow for attenuation of swell noise and vibration noise induced in the streamers while towing them in the ocean (Martin et al., 2000). Moreover, major shifts in the acquisition configurations have enabled the recording of data richer in low frequencies. By towing streamers deeper while at the same time ensuring that data are also recorded in the parts of the spectrum where sea-surface ghosts would normally be introduced, superior signal-to-noise ratio throughout a wide-frequency band can be achieved (e.g., Kragh et al., 2010). Such recording configurations include over/under acquisition and related streamer acquisition (DISCover®) and slanted streamer acquisition (e.g., BroadSeis® or ObliQ®). Finally, dual-component (GeoStreamer®) and multicomponent streamers (Nessie-6®) record both the pressure and the vertical component of particle velocity. Similarly to seabed seismic data, such towed streamer data allow the removal of the sea-surface ghost to increase the temporal bandwidth and to remove the source and receiver ghost images of the subsurface reflectors, thus enabling a much clearer interpretation. On the source side, over/under source configurations and a better understanding of the physics of air guns have enabled improved focusing of emitted energy from air gun sources in the low-frequency part of the spectrum (below 10 Hz; Hopperstad et al., 2012).

A second major shift in marine seismic acquisition configurations was driven by the need for improved illumination of complex subsurface structures (e.g., with subsalt imaging objectives). In combination with the evolution of wave-equation-based imaging algorithms (such as OWEM and RTM), data acquired in various wide and RAZ acquisition configurations, involving one or several vessels (e.g., RAZ, WAZ, and coiled shooting), have resulted in superior subsalt images (e.g., Vigh et al., 2011). Such data are expensive to acquire, requiring a large number of shots, sometimes reoccupying spatial locations where shots have already been acquired but with streamers in different locations and azimuths. Evolving techniques for simultaneous source (or blended) acquisition promise the acquisition of such data more cost-effectively (Beasley, 2008).

Thirdly, marine seismic data are notoriously spatially undersampled, particularly in the crossline receiver direction. New multicomponent streamers and processing techniques (Nessie-6® and IsoMetrix®) allow for the reconstruction of deghosted pressure data on a grid that is finely sampled

both in the inline and crossline directions (Özbek et al., 2010; Robertsson et al., 2008). Such data address many long-standing imaging, multiple elimination, and time-lapse repeatability challenges. Spatial undersampling also exists on the source side where similar concepts have recently been proposed but that require significant further development before they can be deployed in the field (Robertsson et al., 2012).

Finally, marine seismic data have always been acquired as sequential records of data corresponding to individual shots. Instead, a blended seismic acquisition and processing approach would remove this requirement (Berkhout, 2008). Proposed processing approaches of blended seismic data only require that the subsurface has been illuminated through a broad spatial band configuration of sources (and/or source-generated noise such as multiples) and may lead to a major change in the way that we acquire marine seismic data (Berkhout and Verschuur, 2012). However, such techniques are still at a conceptual level but are currently attracting considerable research and development efforts in several research groups in both industry and academia.

Acknowledgments

We are grateful for the excellent reviews of our chapter by Prof. Stewart Greenhalgh, Editor Prof. Lee Slater, and one anonymous reviewer. Schlumberger NExT kindly allowed us to use photos from the course 'Seismic acquisition and processing.'

References

Abma R and Kabir N (2005) Comparisons of interpolation methods. *The Leading Edge* 24: 984–989.

Amundsen L (1993a) Estimation of source array signatures. *Geophysics* 58: 1865–1869.

Amundsen L (1993b) Wavenumber-based filtering of marine point source data. *Geophysics* 58: 1335–1348.

Amundsen L (2001) Elimination of free-surface related multiples without need of the source wavelet. *Geophysics* 66: 327–341.

Amundsen L, Ikelle L, and Berg L (2001) Multidimensional signature deconvolution and free-surface multiple elimination of marine multicomponent ocean-bottom seismic data. *Geophysics* 66: 1594–1604.

Amundsen L, Roesten T, Robertsson JOA, and Kragh E (2005) Rough-sea deghosting of single streamer seismic data using pressure gradient approximations. *Geophysics* 70: V1–V9.

Amundsen L, Westerdahl H, Thompson M, et al. (2010) Multicomponent ocean bottom and vertical cable seismic acquisition for wavefield reconstruction. *Geophysics* 75: W887–W894.

Bagaini C (2006) Overview of simultaneous vibroseis acquisition methods. In: *76th Annual International Meeting, SEG, Expanded Abstracts*, pp. 70–74.

Bale R, Dumitru G, and Probert T (2000) Analysis and stacking of 3D converted wave data in the presence of azimuthal anisotropy. In: *70th Annual International Meeting, SEG, Expanded Abstracts*, pp. 1189–1192.

Barr FJ and Sanders JI (1989) Attenuation of water-column reverberations using pressure and velocity detectors in a water-bottom cable. In: *59th SEG Annual International Meeting, SEG, Expanded Abstracts*, pp. 653–656.

Beasley CJ (2008) Simultaneous Sources: A Technology Whose Time Has Come. In: *78th Annual International Meeting, SEG, Expanded Abstracts*, pp. 2796–2800.

Beasley CJ, Chambers RE, and Jiang Z (1998) A new look at simultaneous sources. In: *68th Annual International Meeting, SEG, Expanded Abstracts*, pp. 133–135.

Beaudoin G and Ross AA (2007) Field design and operation of a novel Deepwater, Wide-Azimuth Node Seismic Survey. *The Leading Edge* 26: 494–503.

Behrens WH (1999) Tuned Bubble Attenuator for Towed Marine Seismic Source. US Patent 5,959,938.

Berg EW, Rykkelid JLE, Woje G, and Svendsen O (2002) Vector fidelity in ocean bottom systems. In: *Offshore Technology Conference*, p. 14114.

Berg E, Svenning B, and Martin J (1995) SUMIC, a new strategic tool for exploration and reservoir mapping. *Journal of Applied Geophysics* 34: 139–140.

Berg EW, Vuillermoz C, Woje G, and Ekmann G (2008) Emerging geophysical technologies: Is planting and re-planting of nodes in a 4C-4D scenario the optimum and most cost-effective solution for field reservoir monitoring? In: *Offshore Technology Conference*, p. 19691.

Berkhout AJ (2008) Changing the mindset in seismic data acquisition. *The Leading Edge* 27: 924–938.

Berkhout AJ and Verschuur DJ (2012) Full wavefield migration – Utilization of multiples in seismic migration. In: *74th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. B041.

Bittleston S, Canter P, Hillesund Ø, and Welker K (2000) Marine seismic cable steering and control. In: *62nd EAGE Conference and Technical Exhibition, Extended Abstracts*, p. L-16.

Bleistein N, Cohen JK, and Stockwell JW (2001) *Mathematics of multidimensional seismic inversion*. New York: Springer.

Breitzke M and Bohlen T (2010) Modelling cumulative sound exposure along a seismic line to assess the risk of seismic research surveys on marine mammals in the Antarctic Treaty Area. In: Popper A and Hawkins A (eds.) *The Effects of Noise on Aquatic Life*, pp. 609–611. New York: Springer, ISBN:978-1-4419-7310-8.

Brice T, Moldoveanu N, and Imamshah A (2012) Multivessel coil surveys in the Gulf of Mexico – Current state of the art and future directions. In: *74th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. X028.

Brummitt JG (1988) Source directivity and its effects on resolution and signature deconvolution. *Exploration Geophysics* 19: 237–240.

Brunellière J, Caprioli P, Grion S, Tilling D, Amundsen L, and Aronsen H (2004) Surface multiple attenuation by up-down wavefield deconvolution: An OBC case study. In: *74th Annual International Meeting, SEG, Expanded Abstracts*, pp. 849–852.

Calvert R (2005a) 4D technology: Where are we, and where are we going? *Geophysical Prospecting* 53: 161–171.

Calvert R (2005b) *Insights and Methods for 4D Reservoir Monitoring and Characterization*. Tulsa, OK: SEG, ISBN:156080128X, SEG/EAGE.

Cambois G, Long A, Parkes GE, Lundsten T, Mattsson A, and Frømyr E (2009) Multilevel air-gun array: A simple and effective way to enhance the low-frequency content of marine seismic data. In: *79th Annual International Meeting, SEG, Expanded Abstracts*, pp. 152–156.

Campbell AJ, Nutt L, Ali S, et al. (2011) An early look at a time-lapse 3D VSP. In: *EAGE Borehole Geophysics Workshop, January, Expanded Abstracts*, p. BG16.

Caprioli PBA, Özdemir AK, Özbek A, et al. (2012) Combination of multi-component streamer pressure and vertical particle velocity – Theory and application to data. In: *74th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. A033.

Ceragioli E, Bovet L, Guilbot J, and Toinet S (2010) The Dalia OBN project. In: *72th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. M021.

Christie PAF, Lunnon ZC, White RS, et al. (2006) iSIMM experience with peak- and bubble-tuned sources for generating low frequencies. In: *68th EAGE Conference and Technical Exhibition, Extended Abstracts, Paper*, p. A034.

Cooper M, Westwater P, Thorogood E, Kristiansen P, and Christie APF (1999) Foinaven active reservoir management towed streamer and buried sea-bed detectors in deep water for 4D seismic. In: *69th Annual International Meeting, SEG, Expanded Abstracts*, pp. 1632–1635.

Cox EJ (2003) *The Source Signature due to the Close Interaction of Marine Seismic Air-Guns*. PhD Thesis, School of Mathematics and Statistics, University of Birmingham.

Cox EJ, Pearson A, Blake JR, and Otto SR (2004) Comparison of methods for modelling the behaviour of bubbles produced by marine seismic air-guns. *Geophysical Prospecting* 52: 451–477.

Day A, Widmaier M, Høy T, and Osnes B (2010) Time-lapse acquisition with a dual-sensor streamer over a conventional baseline survey. *First Break* 28: 79–87.

De Kok R (2012) Directions on ocean-bottom surveying. *The Leading Edge* 31: 415–428.

Dellinger J and Clarke R (2004) Same cable, different vector fidelity: A case study of the Seneca Lake and Valhall LoFS OBC datasets. In: *74th Annual International Meeting, SEG, Expanded Abstracts*, pp. 909–912.

Dragoset WH (1990) Air-gun array specs: A tutorial. *The Leading Edge* 9: 24–32.

Dragoset WH and Jericevic Z (1998) Some remarks on surface multiple attenuation. *Geophysics* 63: 772–789.

Eiken O (2005) Repeatability issues of time-lapse marine seismic data. In: *9th International Congress of the Brazilian Geophysical Society, Salvador, Brazil, Expanded Abstract*.

Eiken O, Haugen GU, Schonewille M, and Duijndam A (2003) A proven method for acquiring highly repeatable towed streamer seismic data. *Geophysics* 68: 1303–1309.

Elboth T, Lilja D, Reif BA, and Andreassen Ø (2010) Investigation of flow and flow noise around a seismic streamer cable. *Geophysics* 75: Q1–Q9.

Elboth T, Reif B, and Andreassen Ø (2009) Flow and swell noise in marine seismic data. *Geophysics* 74: Q17–Q25.

Etgen J, Gray SH, and Zhang Y (2009) An overview of depth imaging in exploration geophysics. *Geophysics* 74: WCA5–WCA17.

Ewing JI and Tirey GB (1961) Seismic profiler. *Journal of Geophysical Research* 66: 2917–2927.

Fayemendy C, Espedal PI, Andersen L, and Lygren M (2012) Time-lapse seismic surveying: A multi-disciplinary tool for reservoir management on Snorre. *First Break* 30: 49–55.

Fokkema JT, Baeten GJM, and Vaage S (1990) Directional deconvolution in the F-X domain. In: *60th Annual International Meeting, SEG, Expanded Abstracts*, pp. 1673–1676.

Forgues E, Meunier J, Hubans C, and Edwards RJ (2006) *Steam Injection Seismic Monitoring Experiment*. CSEG CSPG CWLS Convention.

Fujikawa S and Takahira H (1986) A theoretical study on the interaction between two spherical bubbles and radiated pressure waves in a liquid. *Acta Acustica United with Acustica* 61(3): 188–199.

Gaiser J (2003) Vector-fidelity differences between P-wave first breaks and PS-wave reflections: Implications for compensation of full-azimuth data. In: *73th Annual International Meeting, SEG, Expanded Abstracts*, pp. 822–825.

Giles BF and Johnston RC (1973) System approach to air-gun array design. *Geophysical Prospecting* 21: 77–101.

Gilmore FR (1952) *The growth or collapse of a spherical bubble in a viscous compressible liquid*. *Report No. 26-4*, Pasedena, CA: California Institute of Technology.

Greaves R and Fulp T (1987) Three-dimensional seismic monitoring of an enhanced oil recovery process. *Geophysics* 52: 1175–1187.

Gulunay N (2008) Two different algorithms for seismic interference noise attenuation. *The Leading Edge* 27: 176–181.

Haggerty PE (1954) Method and Apparatus for Cancelling Reverberations in Water Layers. US Patent 2,757,356.

Hampson G, Stefani J, and Herkenhoff F (2008) Acquisition using simultaneous sources. *The Leading Edge* 27: 918–923.

Harlan WS, Langan RT, and Nemeth T (2008) Introduction to the supplement on velocity estimation for depth imaging. *Geophysics* 73: VE1–VE3.

Hatton L and Haddow CM (1991) The effects of source array crossline directivity on 3D migration. *First Break* 9: 427–431.

Hays D, Craft K, Docherty P, and Smit F (2008) An Ocean Bottom Seismic node repeatability study. In: *78th Annual International Meeting, SEG, Expanded Abstracts*, pp. 55–59.

Hegna S and Parkes GE (2011) The low frequency output of marine air-gun arrays. In: *81th Annual International Meeting, SEG, Expanded Abstracts*, pp. 77–81.

Herber R, Weigel W, and Wong HK (1997) Relationship between shot-induced noise and shot interval in seismic refraction experiments at sea. *Marine Geophysical Researches* 19: 257–265.

Herring C (1941) Theory of the pulsations of the gas bubble produced by an underwater explosion. *Underwater Explosion Research*, vol. 2, pp. 35–131. Washington, DC: Office of Naval Research, 1950.

Hill NR (2001) Prestack Gaussian-beam depth migration. *Geophysics* 66: 1240–1250.

Hill D, Combee L, and Bacon J (2006) Over/under acquisition and data processing: The next quantum leap in seismic technology? *First Break* 24: 81–96.

Holbrook WS, Fer I, and Schmitt RW (2009) Images of internal tides near the Norwegian continental slope. *Geophysical Research Letters* 36: L00D10.

Hopperstad JF, Laws RM, and Kragh E (2008c) Where is the center of a multi-depth marine source array? In: *78th Annual International Meeting, SEG, Expanded Abstracts*, pp. 40–44.

Hopperstad JF, Laws RM, and Kragh E (2008a) Fundamental principles of isotropic marine source design. In: *70th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. B025.

Hopperstad JF, Laws RM, and Kragh E (2012) Hypercluster of air-guns – More low frequencies for the same quantity of air. In: *74th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. Z011.

Hopperstad JF, Laws RM, and Melbø AS (2008b) Method for Optimising Acoustic Source Array Performance. US Patent 8,174,927.

Hopperstad JF, Synnevaag JF, and Soppi A (2005) Method and Apparatus for Controlling the Acoustic Output of an Air-Gun. US Patent 7,321,527.

Hopperstad JF, Synnevaag JF, and Vermeer P (2001) An azimuth-invariant source array. In: *71st Annual International Meeting, SEG, Expanded Abstracts*, pp. 52–55.

Huo S, Tsingas C, Kelamis G, Pecholcs PI, and Xu H (2012) Unconstrained simultaneous source land data processing. *Geophysical Prospecting* 60: 608–617.

Ikelle L (2004) The phenomenon of low and high tides in the demultiple process. In: *74th Annual International Meeting, SEG, Expanded Abstracts*, pp. 1858–1860.

Jack I (2001) The coming of age for 4D seismic. *First Break* 19: 24–28.

Keller JB and Kolodner II (1956) Damping of underwater explosions bubbles. *Journal of Applied Physics* 27: 1152–1161.

Kinsler LE, Frey AR, Coppens AB, and Sanders JV (1982) *Fundamentals of Acoustics*. New York: Wiley, ISBN:0-471-09410-2.

Kirk RE, Whitmarsh RB, and Langford JJ (1982) A three-component ocean bottom seismograph for controlled source and earthquake seismology. *Marine Geophysical Researches* 5: 327–341.

Kirkwood JG and Bethe HA (1942) The Pressure wave produced by an underwater explosion I. *Office of Scientific Research and Development Report 588*.

Kragh E and Christie P (2002) Seismic repeatability, normalized RMS and predictability. *The Leading Edge* 21: 640–647.

Kragh E and Laws R (2006) Rough seas and statistical deconvolution. *Geophysical Prospecting* 54: 475–485.

Kragh E, Laws RM, and Hopperstad JF (2012) Reducing the size of a seismic source with a 4C towed streamer. In: *74th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. Z014.

Kragh E, Muyzert E, Curtis T, Svendsen M, and Kapadia D (2010) Efficient broadband marine acquisition and processing for improved resolution and deep imaging. *The Leading Edge* 29: 464–469.

Kragh E, Vigner A, Horne S, Robertsson JOA, and Combee L (2004) Quantifying the horizontal vector fidelity of a marine multicomponent acquisition system. In: *74th Annual International Meeting, SEG, Expanded Abstracts*, pp. 905–908.

Krail PM and Shin Y (2005) Deconvolution of a directional marine source. *Geophysics* 55: 1542–1548.

Landrø M (1992) Modelling of GI gun signatures. *Geophysical Prospecting* 40: 721–747.

Landrø M (1999) Repeatability issues of 3-D VSP data. *Geophysics* 64: 1673–1679.

Landrø M (2008) The effect of noise generated by previous shots on seismic reflection data. *Geophysics* 73: Q9–Q17.

Landrø M, Amundsen L, and Barker D (2011) High-frequency signals from air-gun arrays. *Geophysics* 76: Q19–Q27.

Landrø M and Sollie R (1992) Source signature estimation by inversion. *Geophysics* 57: 1633–1640.

Langhammer J and Landrø M (1993) Temperature effects on airgun signatures. *Geophysical Prospecting* 41: 737–750.

Laws RM (1991) *The Interaction of Marine Seismic Sources*. PhD Thesis, Imperial College, London University.

Laws RM (2010) Cetacean hearing-damage zones around a seismic source. In: Popper A and Hawkins A (eds.) *The Effects of Noise on Aquatic Life*, pp. 473–476. New York: Springer, ISBN:978-1-4419-7310-8.

Laws RM, Hatton L, and Haartsen M (1990) Computer modelling of clustered air-guns. *First Break* 8: 331–338.

Laws RM and Kragh E (2002) Rough seas and time-lapse seismic. *Geophysical Prospecting* 50: 195–208.

Laws RM and Kragh E (2006) Sea surface shape derivation above the seismic streamer. *Geophysical Prospecting* 54: 817–828.

Laws RM and Kragh E (2009) Using Source Elevations to Remove Rough Sea Perturbations. US Patent 0,073,804.

Laws RM, Kragh E, and Morgan G (2008) Are seismic sources too loud? In: *70th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. B026.

Laws RM, Landrø M, and Amundsen L (1998) An experimental comparison of three direct methods of marine source signature estimation. *Geophysical Prospecting* 46: 353–389.

Laws RM and Morice SP (2007) Method of Seismic Surveying, a Marine Vibrator Arrangement, and a Method of Calculating the Depths of Seismic Sources. US Patent 7,257,049.

Laws RM, Parkes GE, and Hatton L (1988) Energy-interaction: The long range interaction of seismic sources. *Geophysical Prospecting* 36: 333–348 (see also the erratum (1990) *Geophysical Prospecting* 38: 104).

Li X-P, Hodges E, Widmaier M, Sundvor E, and Eidsvig S (2004) Seismic repeatability analysis from Oseberg 4D processing – A case history. In: *66th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. P203.

Linden DA (1959) A discussion of sampling theorems. *Proceedings of the Institute of Radio Engineers* 47: 1219–1226.

Long AS (2004) The revolution in seismic resolution: High density 3D spatial sampling developments and results. In: *17th ASEG Geophysical Conference and Exhibition, Extended Abstracts*.

Long A, Mellors D, Allen T, and McIntyre A (2008) A calibrated dual-sensor streamer investigation of deep target signal resolution and penetration on the NW Shelf of Australia. In: *78th Annual International Meeting, SEG, Expanded Abstracts*, pp. 428–432.
Loveridge M, Parkes G, Hatton L, and Worthington M (1984) Effects of marine source array directivity on seismic data and source signature deconvolution. *First Break* 2: 1–23.
Lumley DE (2001) Time-lapse seismic reservoir monitoring. *Geophysics* 66: 50–53.
Lunnon ZC, Christie PAF, and White RS (2003) An evaluation of peak and bubble tuning in sub-basalt seismology: Modelling and results from OBS data. *First Break* 21: 51–56.
Mahdad A, Doulgeris P, and Blacquiere G (2011) Separation of blended data by iterative estimation and subtraction of blending interference noise. *Geophysics* 76: Q9–Q17.
Martell MB, Rothstein JP, and Perot JB (2010) An analysis of superhydrophobic turbulent drag reduction mechanisms using direct numerical simulations. *Physics of Fluids* 22: 065102.
Martin MA and Davis TL (1987) Shear wave birefringence: A new tool for evaluating fractured reservoirs. *The Leading Edge* 6: 22–28.
Martin J, Ozbek A, Combee L, Lunde N, Bittleston S, and Kragh E (2000) Acquisition of marine point receiver seismic data with a towed streamer. In: *70th Annual International Meeting, SEG, Expanded Abstracts*, pp. 37–40.
Meunier J, Huguet F, and Meynier P (2001) Reservoir monitoring using permanent sources and vertical receiver antennae: The Céré-la-Ronde case study. *The Leading Edge* 20: 622–629.
Michell S, Shoshitaishvili E, Chergotis D, Sharp J, and Etgen J (2006) Wide azimuth streamer imaging of Mad Dog: Have we solved the Subsalt imaging problem? In: *76th Annual International Meeting, SEG, Expanded Abstracts*, pp. 2905–2909.
Moldoveanu N (2000) Vertical source array in marine seismic exploration. In: *70th Annual International Meeting, SEG, Expanded Abstracts*, pp. 53–56.
Moldoveanu N (2008) Circular geometry for wide-azimuth towed streamer acquisition. In: *70th EAGE Conference and Technical Exhibition, Expanded Abstracts*, p. G011.
Moldoveanu N (2012) A singular value decomposition algorithm for the attenuation of high energy towed-streamer noise. *First Break* 29: 101–108.
Moldoveanu N, Combee L, Egan M, Hampson G, Sydora L, and Abriel W (2007) Over/under towed-streamer acquisition: A method to extend seismic bandwidth to both higher and lower frequencies. *The Leading Edge* 26: 41–58.
Moore I, Dragoset B, Ommundsen T, Wilson D, Ward C, and Eke D (2008) Simultaneous source separation using dithered sources. In: *78th Annual International Meeting, SEG, Expanded Abstracts*, pp. 2806–2810.
Muijs R, Robertsson JOA, and Holliger K (2007a) Prestack depth migration of primary and surface-related multiple reflections: Part I – Imaging. *Geophysics* 72: S59–S69.
Muijs R, Robertsson JOA, and Holliger K (2007b) Prestack depth migration of primary and surface-related multiple reflections: Part II – Identification and removal of residual multiples. *Geophysics* 72: S71–S76.
Müller C, Woelz S, Ersoy Y, et al. (2009) Ultra-high-resolution marine 2D-3D seismic investigation of the Liman Tepe/Karantina Island archaeological site (Urla/Turkey). *Journal of Applied Geophysics* 68: 124–134.
Olofsson B, Probert T, Kommedal JH, and Barkved OI (2003) Azimuthal anisotropy from the Valhall 4C 3D survey. *The Leading Edge* 22: 1228–1235.
Özbek A (2000) Adaptive beamforming with generalized linear constraints. In: *70th Annual International Meeting, SEG, Expanded Abstracts*, pp. 2081–2084.
Özbek A, Vassallo M, Özdemir K, van Manen DJ, and Eggenberger K (2010) Crossline wavefield reconstruction from multicomponent streamer data: Part 2 – Joint interpolation and 3D up/down separation by generalized matching pursuit. *Geophysics* 75: WB69–WB85.
Parkes GE and Hatton L (1986) *The Marine Seismic Source*. Dordrecht: D. Reidel Publishing, ISBN:90-277-2228-5.
Parkes GE, Ziolkowski AM, Hatton L, and Haugland TA (1984) The signature of an air gun array: Computation from near-field measurements including interactions – Practical considerations. *Geophysics* 49: 105–111.
Parrish JF (1985) Air gun warmup. In: *55th Annual International Meeting, SEG, Expanded Abstracts*, pp. 429–430.
Parrish JF (1986) Air gun warmup pressure and temperature measurements. In: *56th Annual International Meeting, SEG, Expanded Abstracts*, pp. 379–381.
Pascouet A (1991) Something new under the water: The bubbleless air gun. *The Leading Edge* 10: 79–81.
Pettenati-Auziere C, Debouvry C, and Berg E (1997) Node-based sea-bottom seismic, a new way to reservoir management. In: *67th Annual International Meeting, SEG, Expanded Abstracts*, pp. 71–74.
Plesset MS (1949) The dynamics of cavitation bubbles. *Journal of Applied Mechanics* 16(3): 227–282.
Plesset MS and Zwick SA (1952) A non-steady heat diffusion problem with spherical symmetry. *Journal of Applied Physics* 23: 95–98.
Pratt RG and Brenders D (2004) Frequency domain waveform tomography using refracted arrivals. In: *66th EAGE Conference and Technical Exhibition, Extended Abstracts*, p. W8.
Prosperetti A (1984) Bubble phenomena in sound fields: Part 1. *Ultrasonics* 22: 69–77.
Ray CH and Moore NA (1982) High Resolution Marine Seismic Stratigraphic System. PCT Patent W082/00365.
Rayleigh L (1917) On the pressure developed in a liquid during the collapse of a spherical cavity. *Philosophical Magazine* 34: 94–98.
Regone C (2007) Using 3D finite-difference modeling to design wide-azimuth surveys for improved subsalt imaging. *Geophysics* 72: SM231–SM239.
Richardson J (2011) Environmental Assessment of marine vibroseis. *LGL Report TA4604-1* (prepared for the Joint Industry Programme, E&P Sound and Marine Life Contract 22 07-12.
Richardson WJ, Greene CR, Malme CI, and Thomson DH (1995) *Marine Mammals and Noise*. San Diego, CA: Academic Press, ISBN:0-12-588441-9.
Roberts GA and Goulty NR (1988) Directional deconvolution of the seismic source signature combined with prestack migration. *First Break* 6: 247–253.
Roberts GA and Goulty NR (1990) Directional deconvolution of marine seismic reflection data: North Sea example. *Geophysical Prospecting* 38: 881–888.
Robertsson JOA, Halliday DF, van Manen DJ, et al. (2012) Full-wavefield, towed-marine seismic acquisition and applications. In: *74th EAGE Conference and Technical Exhibition, Extended Abstract*, p. Z015.
Robertsson JOA and Kostov C (2010) Multiple Attenuation Method. US Patent 7,710,821.
Robertsson JOA and Kragh E (2002) Rough sea deghosting using a single streamer and a pressure gradient approximation. *Geophysics* 67: 2005–2011.
Robertsson JOA, Moore I, Vassallo M, Özdemir AK, van Manen DJ, and Özbek A (2008) On the use of multicomponent streamer recordings for reconstruction of pressure wavefields in the crossline direction. *Geophysics* 73: A45–A49.
Robertsson JOA, van Manen DJ, Halliday DF, and Laws R (2006) Seismic Data Acquisition and Source-Side Derivatives Generation and Application. US Patent 7,876,642.
Saunders M, Howe D, Kommedal J, and HattonbL Flynn D (2004a) Source signature repeatability and source array directivity – Measurements from the Valhall 4D survey. In: *74th Annual International Meeting, SEG, Expanded Abstracts*, pp. 37–40.
Saunders M, Howe D, Kommedal J, Karran J, and Clayton D (2004b) An efficient, compact omni-directional air gun array for the Valhall 4D/4C seismic survey. In: *66th EAGE Conference and Technical Exhibition, Extended Abstract*, p. H018.
Schalkwijk K, Wapenaar CPA, and Verschuur DJ (2003) Adaptive decomposition of multicomponent ocean-bottom seismic data into downgoing and upgoing P- and S-waves. *Geophysics* 68: 1091–1102.
Schulze-Gatterman R (1972) Physical aspects of the 'airpulser' as a seismic energy source. *Geophysical Prospecting* 20: 155–192.
Shannon CE (1949) Communication in the presence of noise. *Proceedings of the Institute of Radio Engineers* 37: 10–21.
Sheriff RE and Geldart LP (1995) In: *Exploration Seismology*. Cambridge: Cambridge University Press.
Singh SC, Carton H, Tapponnier P, et al. (2008) Seismic evidence for broken oceanic crust in the 2004 Sumatra earthquake epicentral region. *Nature Geoscience* 1: 777–781.
Sirgue L, Denel B, and Gao F (2011) Integrating 3D full waveform inversion into depth imaging projects. In: *81th Annual International Meeting, SEG, Expanded Abstracts*, pp. 2354–2358.
Socco VL, Boiero D, Maraschini M, et al. (2011) On the use of the Norwegian Geotechnical Institute's prototype seabed-coupled shear wave vibrator for shallow soil characterization – II. Joint inversion of multimodal Love and Scholte surface waves. *Geophysical Journal International* 185(1): 237–252.
Söllner W, Brox E, Widmaier M, and Vaage S (2007) Surface-related multiple suppression in dual sensor towed-streamer data. In: *77th Annual International Meeting, SEG, Expanded Abstracts*, pp. 2540–2544.
Sonneland L, Berg L, Eidsvig P, Haugen A, Fotland B, and Vestby I (1986) 2D deghosting using vertical receiver arrays. In: *56th Annual International Meeting, SEG, Expanded Abstracts*, pp. 516–519.
Soubaras R and Dowle R (2010) Variable-depth streamer – A broadband marine solution. *First Break* 28: 89–96.
Southall BL, Bowles AE, Ellison WT, Finneran JJ, Gentry RL, and Greene CR Jr. (2007) Marine mammal noise exposure criteria: Initial scientific recommendations. *Aquatic Mammals* 33: 411–522.

Spetzler J, Xue Z, Saito H, and Nishizawa O (2008) Case story: Time-lapse seismic crosswell monitoring of CO_2 injected in an onshore sandstone aquifer. *Geophysical Journal International* 172: 214–225.
Spitz S (1991) Seismic trace interpolation in the F-X domain. *Geophysics* 56: 785–794.
Stewart J, Shatilo A, Jing C, et al. (2007) A comparison of streamer and OBC seismic data at Beryl Alpha field, U. K. North Sea. *Geophysics* 72: B69–B80.
Stork C, Kapoor J, Zhao W, Dragoset B, and Dingwall K (2006) Predicting and removing complex 3D surface multiples with WEM modeling – An alternative to 3D SRME for wide azimuth surveys? In: *76th Annual International Meeting, SEG, Expanded Abstracts*, pp. 2679–2683.
Strandenes S and Vaage S (1992) Signatures from clustered air-guns. *First Break* 10: 305–312.
Strobbia C, Zarkhidze A, May R, Quigley J, and Bilsby P (2011) Model-based coherent noise attenuation for complex dispersive waves. In: *81th Annual International Meeting, SEG, Expanded Abstracts*, pp. 3571–3575.
Tarantola A (1984) Inversion of seismic reflection data in the acoustic approximation. *Geophysics* 49: 1259–1266.
Taylor P, Hill D, Kapoor J, and Moldoveanu N (2009) Full azimuth: Realizing the benefits. *First Break* 27: 67–71.
Tenghamn R and Dhelie PE (2009) GeoStreamer – Increasing the signal-to-noise ratio using a dual-sensor towed streamer. *First Break* 27: 45–51.
Tenghamn R, Vaage S, and Borreson C (2007) A dual-sensor, towed marine streamer: Its viable implementation and initial results. In: *77th Annual International Meeting, SEG, Expanded Abstracts*, pp. 989–993.
Trad D (2009) Five-dimensional interpolation: Recovering from acquisition constraints. *Geophysics* 74: V123–V132.
Vaage S, Haugland K, and Utheim T (1983) Signatures from single air-guns. *Geophysical Prospecting* 31: 87–97.
Vaage S, Ursin B, and Haugland K (1984) Interaction between air-guns. *Geophysical Prospecting* 32: 676–689.
Vassallo M, Özbek A, Özdemir K, and Eggenberger K (2010) Crossline wavefield reconstruction from multicomponent streamer data: Part 1 – Multichannel interpolation by matching pursuit (MIMAP) using pressure and its crossline gradient. *Geophysics* 75: W853–W867.
Veire HH and Landrø M (2006) Simultaneous inversion of PP and PS seismic data. *Geophysics* 71: R1–R10.
Verschuur DJ, Berkhout AJ, and Wapenaar CPA (1992) Adaptive surface-related multiple elimination. *Geophysics* 57: 1166–1177.
VerWest BJ and Lin D (2007) Modeling the impact of wide-azimuth acquisition on subsalt imaging. *Geophysics* 72: SM241–SM250.
Vigh D, Kapoor J, Moldoveanu N, and Li H (2011) Breakthrough acquisition and technologies for subsalt imaging. *Geophysics* 76: WB41–WB51.
Virieux J and Operto S (2009) An overview of full-waveform inversion in exploration geophysics. *Geophysics* 74: WCC1–WCC26.
Wapenaar CPA, Hermann P, Verschuur DJ, and Berkhout AJ (1990) Decomposition of multicomponent seismic data into primary P- and S-wave responses. *Geophysical Prospecting* 38: 633–662.
Wapenaar CPA, van der Neut J, and Thorbecke J (2012) Deblending by direct inversion. *Geophysics* 77: 9–12.
White RS, Christie APFC, Kuznir NJ, et al. (2002) iSIMM pushes frontiers of marine seismic acquisition. *First Break* 20: 782–786.
Willis HF (1941) Underwater explosions. *British Report WA-47-21*.
Woje G, Berg EW, Rykkelid JI, Svendsen O, and Aanensen O (2002) Vector fidelity analysis of seabed seismic data. In: *64th EAGE Conference and Technical Exhibition, Extended Abstract*, p. E03.
Woodward MJ, Nichols D, Zdraveva O, Whitfield P, and Johns T (2008) A decade of tomography. *Geophysics* 73: VE5–VE11.
Yilmaz Ö (2001) *Seismic Data analysis: Processing, Inversion, and Interpretation of Seismic Data*. Tulsa, OK: Society of Exploration Geophysicists, ISBN:1-56080-098-4.
Ziolkowski AM (1970) A method for calculating the output pressure waveform from an air-gun. *Geophysical Journal of the Royal Astronomical Society* 21: 137–161.
Ziolkowski AM (1980) Source array scaling for wavelet deconvolution. *Geophysical Prospecting* 28: 902–918.
Ziolkowski AM (1984) *Deconvolution*. Dordrecht: D. Reidel Publishing, ISBN:0-934634-62-9.
Ziolkowski AM and Lopes L (1993) Directional deconvolution of marine seismic data in frequency-wavenumber domain. In: *55th EAGE Conference and Technical Exhibition, Extended Abstract*, p. P069.
Ziolkowski AM and Metselaar G (1984) The pressure wavefield of an air gun array. In: *54th Annual International Meeting, SEG Expanded Abstracts*, pp. 274–276.
Ziolkowski AM, Parkes GE, Hatton L, and Haugland TA (1982) The signature of an air-gun array: Computation from near-field measurements including interactions. *Geophysics* 47: 1413–1421.

11.07 Tools and Techniques: Ground-Penetrating Radar

J van der Kruk, Agrosphere, Institute of Bio and Geosciences, Jülich, Germany

© 2015 Elsevier B.V. All rights reserved.

11.07.1	**Introduction to Ground-Penetrating Radar**	209
11.07.2	**EM Wave Propagation: Maxwell's Equations**	210
11.07.2.1	EM Wave Propagation in Homogeneous Space	211
11.07.2.2	EM Wave Propagation in a Homogeneous Half-Space	212
11.07.2.3	EM Wave Propagation in a Horizontally Layered Earth	214
11.07.2.4	EM Wave Propagation in Shallow Subsurface Waveguides	215
11.07.3	**Acquisition Setup of a GPR Survey**	216
11.07.4	**Processing, Imaging, and Inversion of GPR Data**	217
11.07.4.1	Basic Processing Steps	217
11.07.4.2	CMP Processing: Semblance Analysis and Waveguide Dispersion Inversion	217
11.07.4.3	Common-Offset Processing	221
11.07.4.3.1	Polarization information extraction	222
11.07.4.3.2	Imaging/migration	222
11.07.4.4	Ray-Based Tomographic Inversion	226
11.07.4.5	Full-Waveform Inversion	226
11.07.5	**Conclusions**	229
References		230

11.07.1 Introduction to Ground-Penetrating Radar

Ground-penetrating radar (GPR) enables the noninvasive characterization of the subsurface using electromagnetic (EM) waves. Early investigations include measurements of glacier, ice sheet, and permafrost. Nowadays, a multitude of applications exist for this high-resolution geophysical tool, which includes agricultural, archaeological, environmental, geotechnical, hydrologic, and sedimentologic applications. Several books (e.g., Conyers and Goodman, 1997; Daniels, 2004; Jol, 2009), book chapters (e.g., Annan, 2005a,b; Blindow, 2006), and review papers (e.g., Annan, 2002; Davis and Annan, 1989; Huisman et al., 2003; Knight, 2001; Neal, 2004; Peters et al., 1994; Slob et al., 2010) describe the principles of GPR technology, discuss the required processing steps, and show a multitude of existing and emerging applications.

GPR measurements can be applied using different setups; air-launched, crosshole, and surface-based (see **Figure 1**). Here, the source and receiver antennae are indicated as solid arrows, which reflect the dipole nature and indicate that polarized electromagnetic fields are emitted and received respectively. An example of a surface-based GPR section measured across a peat-filled extinct volcano, Durres Maar, is shown in **Figure 2**, where the measured reflections at every position are plotted as function of time. The positive and negative amplitudes of the received pulses are indicated by white and black, respectively. The most striking reflection present in these data is the deepest curved reflection that indicates the basal peat/mineral sediment reflection that can be easily resolved across the peat basin in both directions. Assuming the velocity of the EM waves in peat, $v = 0.035$ m/ns, the time axis can be converted into depth and a maximum reflection depth of approximately 12 m is obtained. Due to the low conductivity in the peat that contains mainly freshwater, a large penetration depth can be obtained. Another clear reflection can be observed at about 1 m deep across the entire profile, which is probably related to a peat harvest. Zooming in on one of these reflections will look similar as the pulses plotted in **Figure 3(a)**.

Using the first arrival times or first-cycle amplitudes, such as indicated in **Figure 3**, enables the estimation of approximate medium properties using ray-based techniques. Conventional processing and inversion methods for GPR data are often using ray-based methods originally developed for seismic data or using far-field high-frequency approximations. In several cases, this approach is sufficient, as long as the main focus is on the propagating nature of the waves. If the focus is on specific EM properties of a material that influence the wave dynamics, such as dipolar vectorial radiation patterns, polarization effects, or reflection coefficients, the scalar seismic techniques are no longer appropriate and specific EM processing and inversion techniques are necessary. For example, the elongated wave trains that are present for receiver depths between 5 and 6 m are due to the presence of a low-velocity waveguide and can only be explained and inverted by using algorithms that are based on Maxwell's equations such as the full-waveform inversion (FWI) that uses all the full-waveform information as indicated in the blue circle.

In the following, the Maxwell's equations are discussed for the EM wave propagation in a homogeneous space, a homogeneous half-space, and a horizontally layered earth. The exact-field radiation patterns due to a dipole source in a homogeneous space and a homogeneous half-space are compared with high-frequency far-field approximations. Next, EM waves are described by transverse electric (TE) and transverse magnetic (TM) modes of propagation, each having corresponding reflection and transmission coefficients. Different acquisition setups are discussed and the special case of wave propagation in a shallow horizontal waveguide layer is analyzed. After discussing the basic processing steps of GPR data

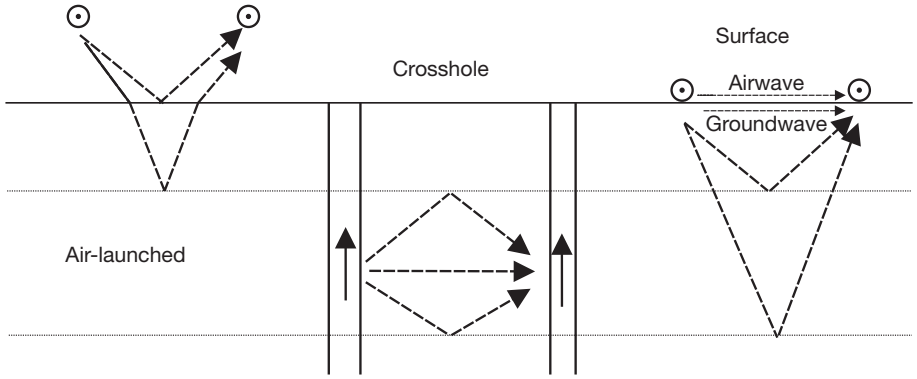


Figure 1 The three most commonly used ground-penetrating radar (GPR) source–receiver setups: air-launched, crosshole, and surface GPR. The orientation of the dipole source and receiver antennae is indicated by the symbol ⊙ for dipoles perpendicular and an arrow for dipoles parallel to the plane, respectively.

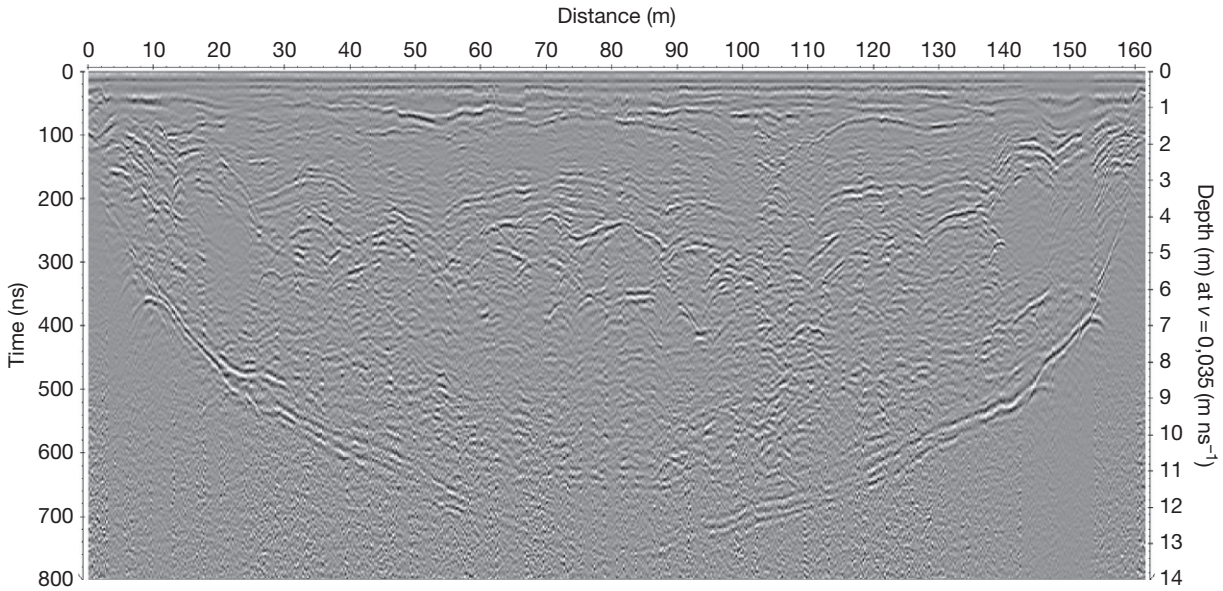


Figure 2 Common-offset surface GPR data measured across a peat-filled extinct volcano.

for common-midpoint (CMP) and common-offset data that are based on scalar approximations, several recent case studies will be described that use new approaches that go beyond traditional analyses of GPR data and explicitly take into account the vectorial EM wave propagation. In this way, waveguide properties can be inverted using a dispersion inversion, the polarization information of a buried scatterer can be extracted using multi-component data, true amplitudes can be obtained from reflectors by eliminating the exact field, and unprecedented high-resolution images can be obtained from the subsurface by using the full waveform in combination with a high-resolution modeling tool.

11.07.2 EM Wave Propagation: Maxwell's Equations

To specify position in a three-dimensional space, the coordinates (x_1, x_2, x_3) are employed in a Cartesian reference frame consisting of three orthogonal base vectors $(\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{i}_3)$, where $\boldsymbol{i}_3$ is pointing downward. Any repeated subscript implies the application of the summation convention. In this way, the position in our configuration may be written as $\mathbf{x}=x_1\ \boldsymbol{i}_1+x_2\ \boldsymbol{i}_2+x_3\ \boldsymbol{i}_3$ or $\mathbf{x}=x_p\boldsymbol{i}_p$ with $p\in\{1,2,3\}$. This subscript notation also enables us to indicate the direction of the electric field written as E_1, E_2, and E_3 or in subscript notation as E_k, where k can take values 1, 2, and 3 (de Hoop, 1995). The EM field equations, also known as Maxwell's equations, are given by

$$-\in_{kmp}\partial_m H_p+(\sigma+\varepsilon\partial_t)E_k=-J_k \qquad [1]$$

$$\in_{jmr}\partial_m E_r+\mu\partial_t H_j=0 \qquad [2]$$

where E_k and H_j are the electric and magnetic field strengths, respectively. J_k is the volume density of electric current, and ∂_m and ∂_t are the partial differentiation with respect to x_m and t, respectively. $\in_{kmp}$ is a unit tensor of rank three (Levi-Civita tensor), which is defined as

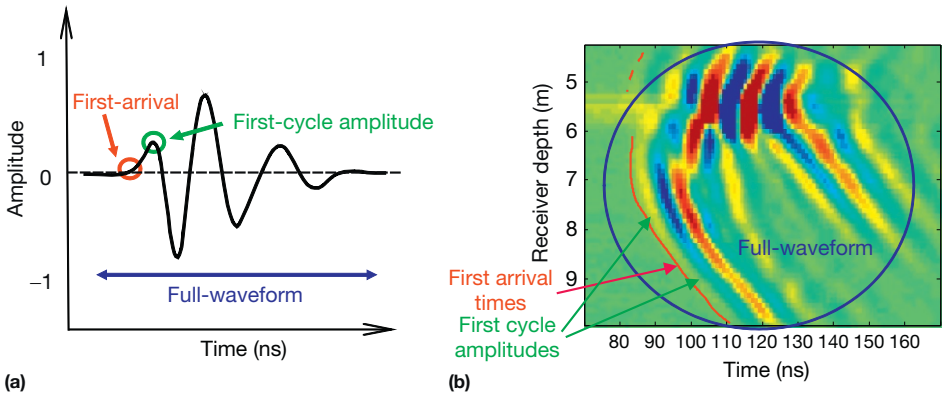


Figure 3 (a) GPR trace amplitude as function of time. (b) Crosshole GPR section measured for a fixed source depth and different receiver depths as function of time. First arrival times and first-cycle amplitude used for ray-based analysis are indicated by red and green arrows, respectively. Ray-based inversion cannot invert the elongated wave trains that are present for receiver depths between 5 and 6 m. Here, the inversion should use the full-waveform information as indicated in the blue circle.

$$\in_{kmp} = \begin{cases} +1 \text{ when } \{k,m,p\} \text{ is an even permutation of } \{1,2,3\}, \\ 0 \text{ when two or more subscripts are equal,} \\ -1 \text{ when } \{k,m,p\} \text{ is an odd permutation of } \{1,2,3\} \end{cases} \quad [3]$$

The permittivity ε can be written as the relative permittivity ε_r times $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m, whereas the permeability μ can be written as the relative permeability μ_r times $\mu_0 = 4\pi \times 10^{-7}$ H/m. In nonmagnetic materials, $\mu_r = 1$. Table 1 shows relative permittivity, conductivity, and velocity values for selected materials, where the velocity v equals $1/\sqrt{\varepsilon\mu} = c_0/\sqrt{\varepsilon_r}$ where the speed of light c_0 is given by $c_0 = 1/\sqrt{\varepsilon_0\mu_0} = 0.299792458\,\text{m/ns}$.

Table 1 Relative permittivity ε_r, conductivity σ, and velocity v for selected materials.

	ε_r	σ *(mS/m)*	v *(m/ns)*
Air	1	0	0.3
Distilled water	80	0.01	0.03
Freshwater	80	0.5	0.03
Seawater	80	3000	0.01
Dry sand	3–5	0.01	0.17–0.13
Saturated sand	20–30	0.1–1	0.07–0.05
Silt	5–30	1–100	0.13–0.05
Clay	5–40	2–100	0.13–0.05
Limestone	4–8	0.5–2	0.15–0.11
Granite	6	0.01–1	0.12
Dry salt	6	0.001–0.1	0.12
Ice	3.18	0.01	0.17

Adapted from Annan AP (2005a) Ground-penetrating radar. In: Butler DK (ed.), *Near Surface Geophysics*. Tulsa, OK: SEG Press, pp. 357–438.

11.07.2.1 EM Wave Propagation in Homogeneous Space

Eliminating the magnetic field H_p by substituting eqn [2] into eqn [1], the general solution for the incident electric field can be written in the frequency domain as

$$\hat{E}_k^i(\mathbf{x}^r) = \int_{x \in D^e} \hat{G}_{kr}^{EJ}(\mathbf{x}^r|\mathbf{x},\omega)\hat{J}_r(\mathbf{x},\omega)\mathrm{d}V \quad [4]$$

where G_{kr}^{EJ} is dyadic Green's function that describes the propagation of the electric field, ω is the angular frequency $2\pi f$, and $\mathrm{d}V$ is the source volume. For a homogeneous space, Maxwell's equations have an analytic solution (de Hoop, 1995) that can be written as

$$\hat{G}_{kr}^{EJ} = \frac{[\partial_k\partial_r - \gamma^2\delta_{kr}]\hat{G}}{\sigma + j\omega\varepsilon} \quad [5]$$

where Green's function $\hat{G}$ satisfies the three-dimensional scalar Helmholtz equation with point-source excitation and is given by

$$\hat{G} = \frac{\exp(-\gamma R)}{4\pi R} \quad [6]$$

and the propagation constant γ is given by

$$\gamma = \sqrt{j\omega\mu(\sigma + j\omega\varepsilon)} \quad [7]$$

and $R = |\mathbf{x}^r\text{-}\mathbf{x}|$. Due to the spatial derivatives, particular direction patterns arise, which depend on the orientation of the source and the direction of observation. Resorting the different contributions shows that the EM field can be divided into three different components: the near field, proportional to R^{-3}; the intermediate field, proportional to R^{-2}; and the far field, proportional to R^{-1} (de Hoop, 1995).

To investigate the vectorial radiation characteristics, the amplitudes of the electric field are analyzed in spherical coordinates (R, φ, θ) in the E-plane $(\varphi = 0)$ and the H-plane $(\varphi = \pi/2)$. The E-plane is parallel to the direction of the current source dipole, while the H-plane is perpendicular to it (see Figure 4). For different angles θ, the electric field is analyzed at a fixed radius of 1 m, which is $\approx 3.3\lambda$ for $f = 500$ MHz, and a relative permittivity of $\varepsilon_r = 4$ and a conductivity $\sigma = 0$ mS/m. This is a representative distance of a possible scatterer with respect to the position of the source and receiver antennae. The amplitude and polarization of the electric total field and far field are plotted in Figure 5(a) and 5(b) for the H-plane and the E-plane, respectively. The circle in the origin of Figure 5(a) indicates the direction of the source, which is perpendicular to the H-plane, and the arrow in the origin of the E-plane in Figure 5(b) indicates that the source direction is parallel to the

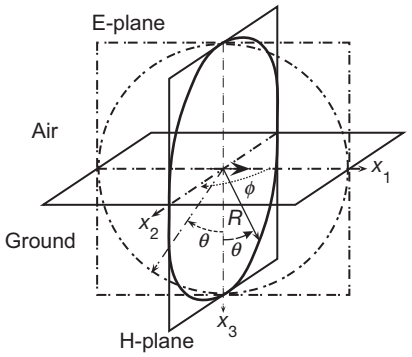


Figure 4 E- and H-planes are parallel and perpendicular to the x_1-directed dipole, respectively.

E-plane. The solid and dashed lines indicate the amplitude of the electric total field and far field, respectively. The electric field in the H-plane is omnidirectional and oriented in the i_2-direction. The electric field component E_θ, which is present in the E-plane, contains both i_1- and i_3-components, whereas a zero is present for $x_3 = 0$. For $\gamma R \gg 1$, the near and intermediate fields have low amplitudes compared to the far field and the contribution of the near field and intermediate field to E_φ and E_θ can be neglected. The difference between the exact field and the far field is about 0.1% and not distinguishable in Figure 5(a) and 5(b). Note that the far field is proportional to $\cos(\theta)$ in the E-plane (de Hoop, 1995).

11.07.2.2 EM Wave Propagation in a Homogeneous Half-Space

To analyze the radiation patterns for a homogeneous half-space, a similar configuration is used as is shown in Figure 4, where a dielectric medium ($\varepsilon_r = 4$) is present for $0 < \theta < \pi/2$, while for $\pi/2 < \theta < \pi$, air is present. Whereas for a homogeneous space, only one velocity is present and one spherical wave is traveling away from the source, there are different contributions in a homogeneous half-space configuration due to the two different velocities of upper and lower half-space and the interface in between. The different contributions of the exact field to the wave fronts in time domain are shown in Figure 6. In addition to the spherical waves in the upper and lower half-space, there are an inhomogeneous wave present in air and a head wave traveling in the ground beyond the critical angle that is given by

$$\theta_{10}^{c} = \sin^{-1}\left(\sqrt{\frac{\varepsilon_0}{\varepsilon_1}}\right) \qquad [8]$$

Closed-form analytic solutions of Maxwell's equations for a horizontal electric dipole present on the interface between two homogeneous half-spaces in space–frequency domain do not exist (Annan, 1973; Radzevicius et al., 2003; Valle et al., 2001). Exact-field solutions are obtained by numerically evaluating the integral equations using standard integration routines (e.g., Valle et al., 2001; van der Kruk, 2001). Far-field high-frequency asymptotic solutions for a lossless half-space are given by Annan (1973), Engheta et al., (1982), and Smith (1984). In general, these far-field expressions can be written as

$$\hat{G}_{kr}^{EJ} = j\omega \hat{A}_{kr} \hat{G} \qquad [9]$$

where $\hat{A}_{kr}$ is an angle-dependent amplitude factor and $j\omega$ is equivalent to a time derivative in time domain.

For low losses, the high-frequency approximation of the propagation constant in eqn [7] can be written as

$$\gamma = j\omega\sqrt{\mu\varepsilon} + \frac{\sigma}{2}\sqrt{\frac{\mu}{\varepsilon}} = \frac{j\omega}{v} + \frac{\sigma}{2}\sqrt{\frac{\mu}{\varepsilon}} \qquad [10]$$

Substitution of eqn [10] in eqn [9] and conversion to time domain (t), where $j\omega$ and $\exp(-j\omega R/v)$ are converted into ∂_t and $\delta(t - R/v)$, respectively, results in the time domain (t) equation for the far-field high-frequency approximation of Green's function:

$$G_{kr}^{EJ} = \frac{\partial_t A_{kr}}{4\pi R} \exp\left(-\frac{\sigma}{2}\sqrt{\frac{\varepsilon}{\mu}} R\right) \delta\left(t - \frac{R}{v}\right) \qquad [11]$$

The far-field expressions have a geometric spreading of $1/R$ and only describe the body wave in air and ground. Note that the velocity v in air and that of the ground are different. Figure 5(c)–5(d) depicts the exact- and far-field spherical electric components E_φ in the H-plane and E_θ in the E-plane at a radius of 1 m $\approx 3.3\lambda$ for $f = 500$ MHz. The far-field contribution has a zero at the interface and a maximum and a minimum at the critical angle θ_c in the H-plane and E-plane, respectively. However, the maximum and minimum responses for the exact-field radiation pattern in the H-plane and E-plane appear at angles larger than the critical angle and are less pronounced than those computed for far-field models. Similar observations have been made in comparison with far-field expressions to exact-field finite-difference time domain (FDTD) modeling results (Holliger and Bergmann, 1998; Radzevicius et al., 2003). Investigating the far-field expressions for the airwave and ground wave that travel directly from source to receiver shows that these contributions are proportional to $1/x_2^2$ (Dai and Young, 1997; Di Matteo et al., 2013) and can be written for the broadside setup (see Fig. 7) as

$$G_{11}^{\text{airwave}} = \frac{Z_0}{2\pi(1-\varepsilon_r)x_2^2} \partial\left(t - \frac{x_2}{c_0}\right) \qquad [12]$$

$$G_{11}^{\text{groundwave}} = \frac{-\sqrt{\varepsilon_r} Z_0}{2\pi(1-\varepsilon_r)x_2^2} \exp\left(-\frac{\sigma}{2} Z_1 x_2\right) \partial\left(t - \frac{x_2}{v}\right) \qquad [13]$$

where $Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}$ and $Z_1 = \frac{Z_0}{\sqrt{\varepsilon_1}}$.

The exact-field expressions describe all wave phenomena including the head wave in the ground and the inhomogeneous wave in air. The large difference between the exact- and the far-field expressions is probably due to the large contribution of the intermediate field near the critical angle θ_c, which describes the head wave. In Figure 5(e) and 5(f), the far field and exact field are compared to a source elevated 0.1λ above the interface. In this case, a much larger portion of the field is radiated into the air (Smith, 1984; Slob and Fokkema, 2002a, 2002b; van der Kruk, 2004). The far-field amplitudes differ strongly from the exact-field ones, not only near the critical angle but also in the vertical below the source (i.e., for $\theta = 0$). Moreover, also a more focused footprint is obtained, which is reducing the angular coverage of measured reflections.

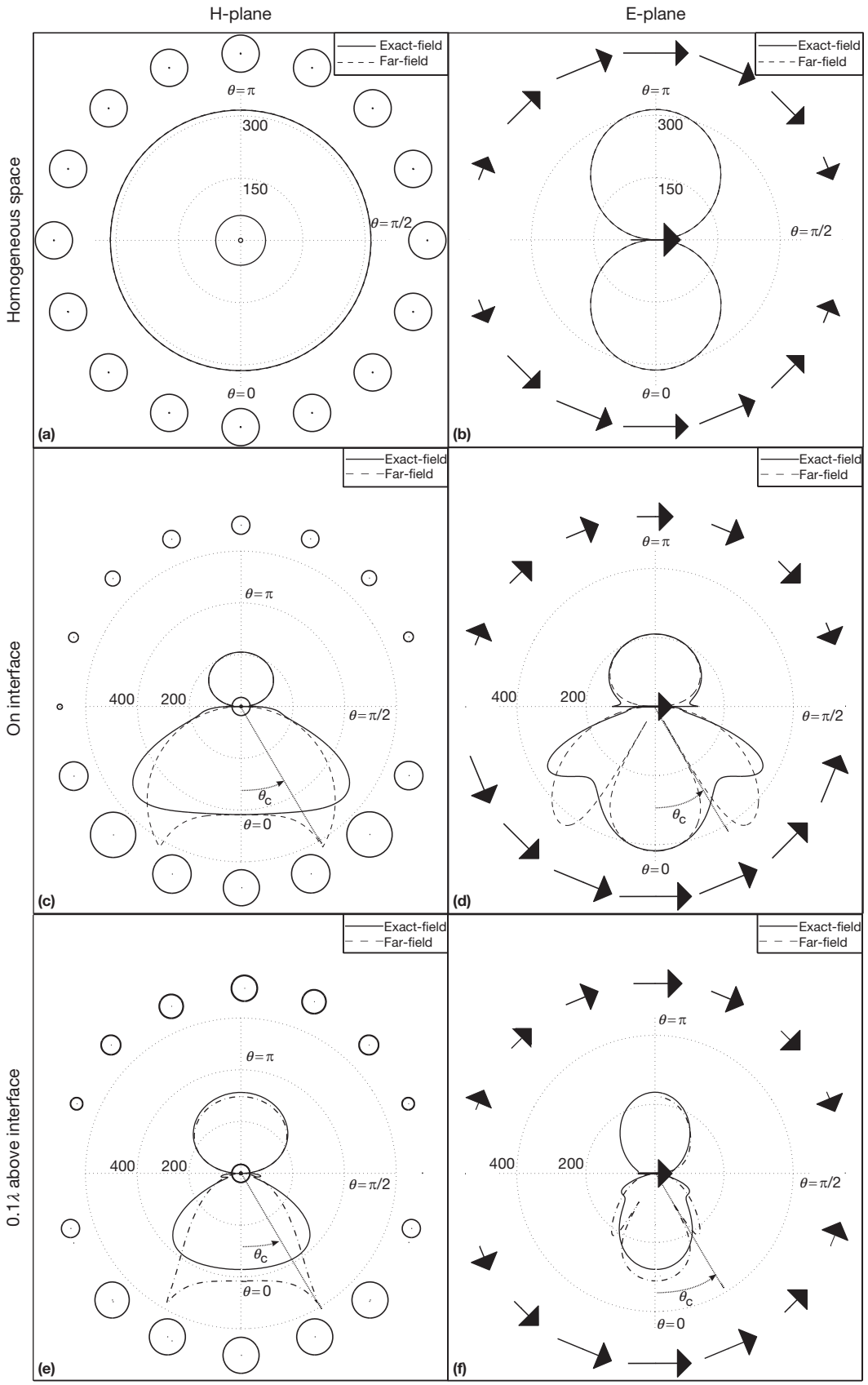


Figure 5 (a) and (b) show E_φ and E_θ exact- (solid) and far-field (dashed) amplitudes in the H-plane and E-plane, respectively, for a homogeneous space; (c) and (d) for a dipole antenna present on the interface between air ($\varepsilon_r = 0$) and the subsurface ($\varepsilon_r = 4$); (e) and (f) for a dipole antenna elevated 0.1λ above the interface between air and the subsurface at a radial distance of $R = 1\ \text{m} \approx 3.3\lambda$, a frequency $f = 500$ MHz, $\varepsilon_r = 4$, and $\sigma = 0\ \text{S m}^{-1}$. The orientation of the source in the center of the H-plane and E-plane is indicated by the symbol $\odot$ and an arrow, respectively. The orientation of the emitted electric field is indicated by circles and arrows surrounding the amplitude patterns, which shows that the field is perpendicular to the H-plane in (a, c, e) and parallel to the E-plane in (b, d, f).

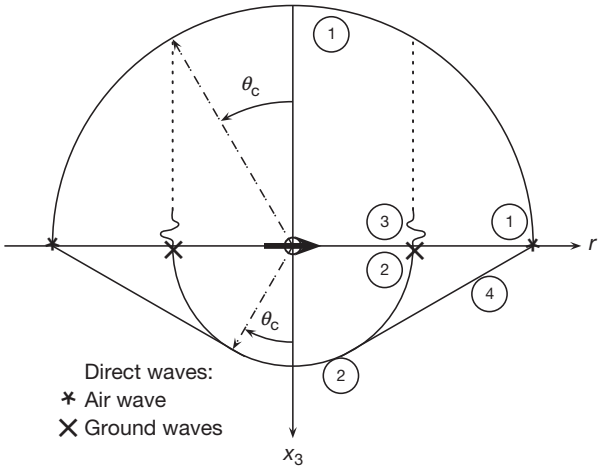


Figure 6 Wave fronts generated by a dipole source on the surface of a half-space earth with 1: body wave in air, 2: body wave in ground, 3: inhomogeneous wave in air, and 4: head wave in ground.

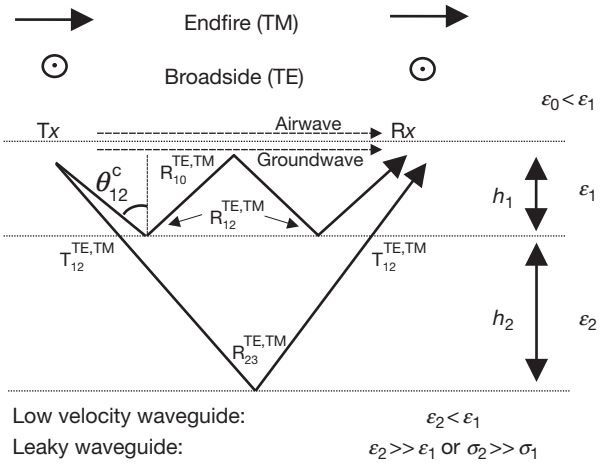


Figure 7 Horizontally layered medium that consists of a shallow waveguide layer, where multiple reflections occur. For the broadside and endfire source (Tx)–receiver (Rx) configurations, the TE and TM reflection and transmission coefficients are indicated.

The results in Figure 5 show that far-field expressions approximate the exact field well within a homogeneous space, which is the case for air-launched and borehole GPR measurement setups. For antennae present on the interface between air and ground, the far-field expressions differ strongly from the exact-field ones.

11.07.2.3 EM Wave Propagation in a Horizontally Layered Earth

When a plane wave with an incident angle θ_I with respect to the normal is reflected at the interface between medium 1 and medium 2 as shown in Figure 7, angle-dependent reflection coefficients occur that are also depending on the orientation of the electric field with respect to the plane of incidence. The TE polarization has the electric field polarized perpendicular to the plane of incidence, whereas the TM polarization has the magnetic field polarized perpendicular to the plane of incidence. The pertaining TE and TM reflection coefficients are given by

$$r_{12}^{\mathrm{TE}}(\theta_i) = \frac{\sqrt{\varepsilon_1}\cos(\theta_i) - \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2(\theta_i)}}{\sqrt{\varepsilon_1}\cos(\theta_i) + \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2(\theta_i)}} \quad [14]$$

$$r_{12}^{\mathrm{TM}}(\theta_i) = \frac{\varepsilon_1\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2(\theta_i)} - \varepsilon_2\cos(\theta_i)}{\varepsilon_1\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2(\theta_i)} + \varepsilon_2\cos(\theta_i)} \quad [15]$$

When the numerator in eqn [15] goes to zero, no reflected waves for the TM mode exist. This angle is referred to as Brewster's angle and is given by

$$\tan(\theta_i) = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} \quad [16]$$

For the special case of normal incidence, $\theta_i = 0$, which is the case for a zero-offset configuration, both TE and TM reflection coefficients return the same reflection coefficient given by

$$r_{12}^{\mathrm{TE}}(0) = r_{12}^{\mathrm{TE}}(0) = \frac{\sqrt{\varepsilon_1} - \sqrt{\varepsilon_2}}{\sqrt{\varepsilon_1} + \sqrt{\varepsilon_2}} \quad [17]$$

Similarly, the TE and TM transmission coefficients can be given by

$$t_{12}^{\mathrm{TE}}(\theta_i) = \frac{2\sqrt{\varepsilon_1}\cos(\theta_i)}{\sqrt{\varepsilon_1}\cos(\theta_i) + \sqrt{\varepsilon_2 - \varepsilon_1 \sin^2(\theta_i)}} \quad [18]$$

$$t_{12}^{\mathrm{TM}}(\theta_i) = \frac{2\varepsilon_1\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2(\theta_i)}}{\varepsilon_1\sqrt{\varepsilon_2 - \varepsilon_1 \sin^2(\theta_i)} + \varepsilon_2\cos(\theta_i)} \quad [19]$$

For the special case of normal incidence, $\theta_i = 0$, which is the case for a zero-offset configuration, both TE and TM reflection coefficients return the same reflection coefficient given by

$$t^{\mathrm{TE}}(0) = t^{\mathrm{TM}}(0) = \frac{2\sqrt{\varepsilon_1}}{\sqrt{\varepsilon_1} + \sqrt{\varepsilon_2}} \quad [20]$$

For a horizontally layered medium as shown in Figure 7, the reflected electric field E_{refl} can be written as

$$E_{\mathrm{refl}}^{\{\mathrm{TE,\,TM}\}} = E_{\mathrm{inc}}^{\{\mathrm{TE,\,TM}\}}\left(r_{01}^{\{\mathrm{TE,\,TM}\}} + t_{01}^{\{\mathrm{TE,\,TM}\}} t_{10}^{\{\mathrm{TE,\,TM}\}} \sum_{n=0}^{\infty}\left(r_{10}^{\{\mathrm{TE,\,TM}\}} r_{12}^{\{\mathrm{TE,\,TM}\}} \beta_1\right)^n\right) \quad [21]$$

where β_1 describes the time delay that is associated with the multiple reflections and

$$\beta_1 = \exp[-2\gamma_1 h_1 \cos(\theta)] \quad [22]$$

Note that the sum in eqn [21] is a geometric continuation of common ratio lower than one and can be rewritten as

$$\sum_{n=0}^{\infty}\left(r_{12}^{\{\mathrm{TE,\,TM}\}} r_{10}^{\{\mathrm{TE,\,TM}\}} \beta_1\right)^n = \frac{1}{1 - r_{12}^{\{\mathrm{TE,\,TM}\}} r_{10}^{\{\mathrm{TE,\,TM}\}} \beta_1} \quad [23]$$

Expressions for the wave propagation in a multilayered medium have been described in a wide range of studies (Annan, 1973; Dai and Young, 1997; Kong, 1986; Michalski and Mosig, 1997).

11.07.2.4 EM Wave Propagation in Shallow Subsurface Waveguides

At locations where a surface layer with a thickness in the order of a wavelength overlies a substrate medium that has a lower permittivity, or a much larger permittivity/conductivity than the surface layer (Figure 7), pronounced dispersion of GPR waves can be observed and the surface layer acts as a waveguide. When the substrate medium has a lower permittivity than the waveguide layer, $\varepsilon_0 < \varepsilon_2 < \varepsilon_1$, complex roots will be present in the expressions for the TE and TM reflection coefficients beyond the critical angle

$$\theta_{12}^c = \sin^{-1}\left(\sqrt{\frac{\varepsilon_2}{\varepsilon_1}}\right) \quad [24]$$

and total reflection occurs beyond the critical angle. Note that also on the upper interface, total reflection occurs (see eqn [8]), such that the layer acts as a low-velocity waveguide. When the lower interface is a strong reflector, $\varepsilon_2 \gg \varepsilon_1$ or $\sigma_2 \gg \sigma_1$, total reflection only occurs at the upper interface, and some energy is still transmitted across the lower interface; hence, it is a leaky waveguide. In both cases, the EM waves are trapped within the waveguide and the radar energy is internally reflected, resulting in a series of multiple reflected waves (multiples) that interfere and manifest themselves as a package of dispersed waves.

The fundamental equation of modal theory is given by Budden (1961) and Arcone (1984)

$$1 - R_{12}^{\{\mathrm{TE,\,TM}\}} R_{10}^{\{\mathrm{TE,\,TM}\}} \beta_1 = 0 \quad [25]$$

where the reflection coefficients are given by eqns [13] and [14] and the propagation constant is given by eqn [7]. Note that here, the denominator in eqn [23] is set equal to zero. To solve this equation, the two reflection coefficients must have unit amplitude, such that the total reflection occurs at both interfaces. For progressively increasing angle θ, a total reflection occurs first at the upper interface and then at the lower interface. Therefore, for mode propagation, the critical angle of the lower boundary must be reached. Note that although the reflection amplitude equals one, the phase still changes as function of the angle. For optimum constructive interference, the total phase change after successive reflections at the upper and lower interfaces must equal $2m\pi$ radians, where m is an integer. Different values of m correspond to different guided-wave modes, whereas the fundamental mode has a value of $m = 0$. Numerically, eqn [25] is solved for models with medium parameters ε_1, ε_2, and h, using a zero-finding routine, and returns a decreasing phase velocity v for increasing frequency f (van der Kruk et al., 2006).

The cutoff frequency is the lowest frequency at which a wave can travel through a waveguide. For these asymmetric waveguides (i.e., material properties on either side of the waveguide are not the same $\varepsilon_2 \neq \varepsilon_0$), the cutoff frequency is given by (van der Kruk, 2006)

$$f_c^{\{\mathrm{TE}_m,\,\mathrm{TM}_m\}} = \frac{\left[2\pi m + \phi^{\{\mathrm{TE}_m,\,\mathrm{TM}_m\}}\left(\theta_{12}^c\right)\right] c_0}{4\pi h\sqrt{\varepsilon_1 - \varepsilon_2}} \quad [26]$$

where

$$\phi^{\mathrm{TE}_m}\left(\theta_{12}^c\right) = \tan^{-1}\left(\frac{2\sqrt{\varepsilon_1 - \varepsilon_2}\sqrt{\varepsilon_2 - \varepsilon_0}}{\varepsilon_0 + \varepsilon_1 - 2\varepsilon_2}\right)$$

$$\phi^{\mathrm{TM}_m}\left(\theta_{12}^c\right) = \tan^{-1}\left(\frac{-2\varepsilon_0\varepsilon_1\sqrt{\varepsilon_1 - \varepsilon_2}\sqrt{\varepsilon_2 - \varepsilon_0}}{-\varepsilon_0^2(\varepsilon_1 - \varepsilon_2) + \varepsilon_1^2(\varepsilon_2 - \varepsilon_0)}\right) + \pi \quad (27)$$

The different phases, ϕ^{TE} and ϕ^{TM}, result in different cutoff frequencies for the TE and TM modes, where $f_c^{\mathrm{TE}_m} < f_c^{\mathrm{TM}_m}$. Note that in contrast to the seismic case for which the cutoff frequencies of the fundamental Rayleigh and Love wave modes are zero (Aki and Richards, 1980), both the TE and the TM fundamental modes have nonzero cutoff frequencies. When the upper and lower homogeneous half-spaces are equal, $\varepsilon_0 = \varepsilon_2$, the critical angles for the two interfaces are equal, and the expression for the cutoff frequency given in [26] reduces to

$$f_c^{\{\mathrm{TE}_m,\,\mathrm{TM}_m\}} = \frac{mc_0}{2h\sqrt{\varepsilon_1 - \varepsilon_2}} \quad [28]$$

For a leaky waveguide where the lower interface is a strong reflector, $R_{12}^{\{\mathrm{TE,TM}\}} = -1$ is substituted in eqn [25] and the following expressions for the cutoff frequencies are obtained (van der Kruk et al., 2007a,b):

$$f_c^{\mathrm{TE}_m} = \frac{(2m+1)c_0}{4h\sqrt{\varepsilon_1 - 1}}$$

$$f_c^{\mathrm{TM}_m} = \frac{2mc_0}{4h\sqrt{\varepsilon_1 - 1}} \quad [29]$$

To show the underlying wave phenomena, three-dimensional FDTD modeling was performed of EM waves propagating through a low-velocity surface waveguide with $\varepsilon_1 = 23$ and $h_1 = 0.35$ m overlying a half-space with $\varepsilon_2 = 11$. A dipole current source (Tx) is present at $(x_1, x_2, x_3) = (0,0,0)$ and emits a Gaussian pulse with a center frequency of 80 MHz. Figure 8 shows for (a) $t = 50$ and (b) $t = 100$ ns the electric field (E_1) distribution in the plane $x_1 = 0$. The body wave traveling in air (including the direct airwave) has already traveled outside of the plotting range. The body wave traveling in the ground (lower half-space) is clearly visible (Figure 8(a)). The head wave (critical refraction from the lower half-space) is also indicated within the waveguide. Within the surface waveguide, the GPR signal does not consist of one single pulse but of a more elongated wave train due to multiple internal reflections within the waveguide. The inset in Figure 8(b) shows the modal electric field distributions, TE_0, and TE_1, that can propagate through a dielectric waveguide due to the constructive interference caused by the multiple internal reflections. The propagating modes depend on the TE reflection coefficients at either interface and the height and permittivity of the waveguide. The TE_0 and TE_1 modes can be clearly identified in Figure 8 within the yellow circles. The fundamental TE_0 mode has a positive (red) or negative (blue) amplitude over the whole vertical range of the waveguide, whereas the higher-order modes contain alternating positive and negative (or vice versa) amplitudes. Note that the overall shape of the electric field distributions (pattern and amplitude in the y-direction within the waveguide) in Figure 8(a) and 8(b) differs because the different frequencies travel with different phase velocities.

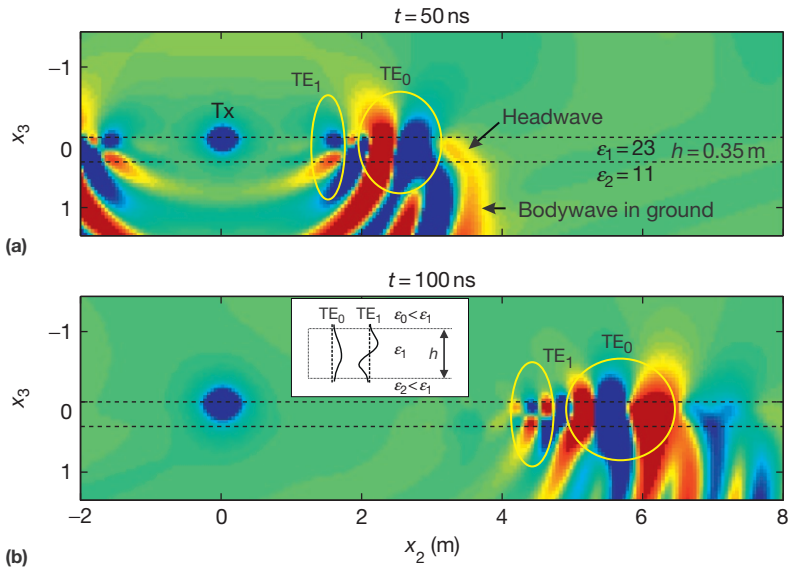


Figure 8 Snapshots of 3-D FDTD modeling of the electric field E_1, emitted by a x_1-oriented dipole present at the origin at times (a) $t = 50$ and (b) 100 ns. High-amplitude elongated wave energy is traveling through the surface waveguide with $\varepsilon_1 = 23$ and $h = 0.35$ m present between the two dashed lines overlying a half-space with $\varepsilon_2 = 11$. The TE_0 and TE_1 modes (see inset figure in (b)) can be clearly identified within the yellow circles. Adapted from van der Kruk J, Jacob RW, and Vereecken H (2009a) Identifying dispersive GPR signals and inverting for surface wave-guide properties. *The Leading Edge* 28: 1234–1239.

11.07.3 Acquisition Setup of a GPR Survey

Conventional GPR systems consist of two antennae: a transmitting dipole antenna, which emits a vectorial EM wave into the subsurface, and a dipole receiving antenna, which detects a specific polarization of the back-traveling EM waves (see Figure 9). Three measurement modes can be used: crosshole, off-ground, and surface GPR (see Figure 1), which have their own drawbacks, benefits, and need for specific processing steps. Crosshole GPR operates between two (sub)parallel boreholes at a certain distance in which the vertical transmitting and receiving antennae are located. Therefore, the emitted energy propagates from the source antenna to the receiver antenna, and the arrivals for a large number of source and receiver positions are used to obtain velocity and attenuation information from between the boreholes. Although crosshole GPR is widely used (e.g., Binley et al., 2001), the method is evidently restricted by the availability of borehole installations. Due to their noninvasive nature, off-ground GPR and surface GPR overcome this limitation and are more often applied. Off-ground GPR uses air-launched antennae installed at a sufficient height above the surface and is often used for soil characterization (Jadoon et al., 2012; Lambot et al., 2004; Redman et al., 2002; Weihermüller et al., 2007) and asphalt or concrete characterization (Hugenschmidt, 2002; Kalogaropoulos et al., 2011; Saarenketo and Scullion, 2000). Since mainly vertical wave propagation occurs, the applied measurement setup often consists of a constant offset between the source and the receiver antennae that are present in their respective H-plane, also called broadside configuration, to optimize the angular coverage (see Figure 5). To interpret and invert the measured reflected waves, horizontal layers are often assumed. The main challenges for off-ground GPR are its sensitivity to surface roughness (Lambot et al., 2006) and the limited

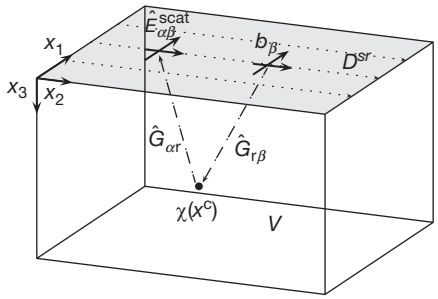


Figure 9 The configuration of the four possible source–receiver setups. Solid arrows indicate antenna orientations, b_β indicates the dipole source orientation, and $\hat{E}^{scat}_{\alpha\beta}$ indicates the electric field measured by a dipole receiver oriented in the x_α-direction. Dashed arrows indicate the EM wave propagation paths from source to scatterer and from point scatterer $\chi(x^c)$ to receiver.

penetration depth of the signal due to the relatively large reflection occurring at the interface between the air and the ground. Compared to antennae that are lying on the interface, lifting the antennae 0.1λ above the interface results in reduced amplitudes that are emitted in the subsurface and increased amplitudes emitted in air (see Figure 5 (e) and (f)). Surface GPR antennae emit more energy into the subsurface (see Figure 5 (c) and (d)) such that higher penetration depth can be obtained. Moreover, wider radiation patterns also enable the reconstruction of subsurface objects with higher spatial resolution. In addition to the common-offset source–receiver setup, a commonly used setup is the wide-angle reflection–refraction (WARR) method (Huisman et al., 2003), where the reflections have a CMP or the source position remains fixed. Whereas the common-offset method enables the mapping over larger areas, the CMP and

WARR methods enable an improved characterization of the subsurface since the different wave types, such as direct airwave and ground wave and reflected and refracted waves, can be clearly distinguished when increasing the distance between the source and the receiver antennae. In addition to the broadside configuration, also the end-fire configuration, where the parallel oriented source and receiver antennae are present in their respective E-plane, contains important information due to the relatively broad radiation patterns in both the H-plane and the E-plane (see **Figure 5(c)** and **5(d)**). **Figure 7** shows the broadside and end-fire configurations for a horizontally layered medium.

11.07.4 Processing, Imaging, and Inversion of GPR Data

To obtain a representative image of the subsurface from the measured GPR data, specific processing steps are needed that are based on the EM wave propagation discussed before. First, some processing steps that are used for most measurements setups are discussed. Then, specific processing steps are discussed and shown for selected experimental CMP, common-offset, and crosshole data.

11.07.4.1 Basic Processing Steps

As shown in **Figure 3**, the measured reflections contain positive and negative peaks that should be plotted on either side of the zero amplitude level. However, due to very low-frequency components that are caused by inductive phenomena or possible instrumentation dynamic range limitations (Annan, 2005a), the measured reflections are often not centered along the zero amplitude level. Simple temporal filtering, also called dewowing, is usually applied to remove these low-frequency components and center the received wavelets along the zero amplitude level. Due to the geometric spreading and the presence of the attenuation due to conductivity losses as described in eqn [11], the received amplitudes decay with increasing distance between the source, reflection point, and receiver. To compensate for this amplitude decay, a time-varying gain is usually applied to enable a visual quality check during acquisition. Different time gains exist such as automatic gain correction, which is widely used to display seismic data (Yilmaz, 2000). Spherical and exponential compensation gain functions correct for the geometric spreading and attenuation due to conductivity losses as described in eqn [9]. Another frequently applied gain function is the inverse envelope gain function that uses the Hilbert transform to calculate the envelope of the traces (Gross et al., 2003). Remaining high-frequency noise is often removed by using a band-pass frequency filter. The data shown in **Figure 2** were obtained after dewow, time-zero correction, energy decay correction, and a band-pass filter.

11.07.4.2 CMP Processing: Semblance Analysis and Waveguide Dispersion Inversion

Common processing steps where the velocity of the EM waves can be estimated use ray-based techniques such as ground wave picking (Galagedara et al., 2003; Steelman and Endres, 2010), reflected wave analysis (Bradford, 2008; Endres et al., 2000; Garambois et al., 2002; Greaves et al., 1996; Jacob and Hermance, 2004; Tillard and Dubois, 1995), refracted wave analysis (Bohidar and Hermance, 2002), or a combined analysis of ground, reflected, and refracted waves (Huisman et al., 2003; van Overmeeren et al., 1997). Most of these techniques have been developed for the processing of seismic data and to determine the seismic wave velocity (Yilmaz, 2000). The advantage of these ray-based methods is that the medium properties can be easily computed from the measured data. However, since only the travel time or phase information of the GPR data is used (see also eqns [9]–[11]), these methods only give reliable values for the permittivity of the subsurface when far-field approximations are valid.

Not only to use GPR as a geometric tool to localize reflectors but also to go one step further and identify the nature of the reflectors and the surrounding media, the amplitude and phase of the TE and TM reflection coefficients can be used (e.g., Hollender and Tillard, 1998). The TE and TM reflection coefficients also contain information about fracture aperture and filling material. However, the relationships with the GPR signal amplitude, phase, and frequency are not well understood. Tsoflias et al. (2004) observed for surface-based multipolarization GPR measurements across vertical fractures a phase lead when the incident electric field component is oriented perpendicular to the plane of the fracture. Using an experimental setup, Tsoflias and Hoch (2006) investigated multipolarization GPR for characterizing thin, vertical fracture properties. They showed that at high angles of incidence, variable polarization GPR signals display characteristic and quantifiable phase and amplitude responses that are related to thin-layer properties. The measured GPR data were in agreement to analytic solutions of plane-wave oblique-incidence transmission through layered media. Tsoflias and Becker (2008) used time-lapse multifrequency GPR to investigate the influence of fracture-fluid salinity and show that for increasing fluid electrical conductivity, lower-frequency signals exhibit greater amplitude and phase-delay changes compared with high-frequency systems.

Amplitude versus offset (AVO) analysis can be applied to GPR data to yield information not available from a traditional common-offset survey. In the specific case of two subsurface anomalies, a silty-sand lens and an Light Non-Aqueous Phase Liquid (LNAPL)-saturated sand lens, the common-offset profiles look similar, whereas the amplitude variations with offset can provide enough information to make quantitative judgments about different physical properties of each anomaly (Baker, 1998). Hollender and Tillard (1998) investigated the reflection coefficient for reflections from a half-space and a thin layer using plane wave propagating in 1-D models and the Jonscher parameterization (Jonscher, 1977). Zeng et al. (2000) investigated the AVO response by varying the Cole–Cole parameters (Cole and Cole, 1941) using a 2.5-D finite-difference solution of Maxwell's equations. Reppert et al. (2000) showed that the dielectric permittivity contrast at a boundary can be estimated using Brewster's angle (see eqn [16]). The conversion of CMP field data into reflection coefficient properties is a delicate operation because many variables influence this relationship, such as source–receiver coupling and radiation patterns, geometric spreading, attenuation, and emitted source wavelet (Bradford

and Deeds, 2006; Deparis and Garambois, 2009). Assuming that the coupling factor does not vary significantly for different offsets, the attenuation does not vary significantly with offset, and normalizing the observed amplitude at any offset to the amplitude at near offset results in the elimination of many influencing factors. Remaining corrections need to be carried out for the radiation patterns, geometric spreading, and attenuation. Semiempirical radiation pattern corrections were introduced by Bradford and Deeds (2006), and an analytic solution of the thin-bed reflectivity was successfully applied to two case studies dealing with NAPL-contaminated zones where the amplitudes of the reflected wave were analyzed by using the envelope of the signal. Deparis and Garambois (2009) used an FDTD modeling tool developed by Lampe and Holliger (2005) to numerically simulate radiation patterns for bow-tie GPR antennae using a Wu and King (1965) profile scaled to one-fourth of the nonreflecting Wu–King load (Streich and van der Kruk, 2007a). An inversion algorithm based on the neighborhood algorithm (Sambridge, 1999a,b) was used to determine the Jonscher parameter (Jonscher, 1977) for the matrix and the fracture of synthetic data and a field data set acquired on a vertical cliff. The frequency dependency of the reflection coefficient was of major importance to obtain a satisfactory and unique solution (Deparis and Garambois, 2009).

In the following, a case study is described where repeated GPR measurements were performed and analyzed by using a semblance analysis and waveguide dispersion inversion (see also van der Kruk et al., 2010b). The field site is located in south eastern New England (Jacob and Hermance, 2004) and consists of glacially stratified drift and outwash deposits. The 200 MHz data were collected by incrementally increasing the source–receiver offset, using a total step size of 0.2 m to a maximum offset of 20 m along a common measurement line. The GPR signal was sampled every 0.8 ns and the data were collected with an internal 64-fold stack per trace. **Figure 10** shows broadside and end-fire CMP data after applying a dewow filter and a time-dependent gain that was based on a spatially averaged Hilbert transform (Gross et al., 2003). **Figure 10(a)** and **10(c)** shows the broadside and end-fire data measured in dry conditions. The airwave and several reflections can be clearly observed and identified for almost the whole offset range. After scaling of the data to yield roughly equal amplitudes and compensating for the normal move out for a range of velocities v, a semblance spectrum (Taner and Koehler, 1969; Yilmaz, 2000) can be calculated for the data f_{ij}, where i and j indicate trace and sample, respectively, using

$$S(t_j, v) = \frac{\sum_{j}^{j+N}\left[\sum_{i=1}^{M} f_{ij}\right]^2}{M\sum_{j}^{j+N}\sum_{i=1}^{M}\left(f_{ij}\right)^2} \qquad [30]$$

where M is the number of traces. To ensure that the entire georadar wavelet is considered in the correlation process, the number of samples N of the semblance tube should be set equal to the inverse of the lowest dominant signal frequency. Large semblance amplitudes within the semblance spectra of **Figure 10(b)** and **10(d)** indicate the root-mean-square velocity of the reflections that can be converted to interval velocity using Dix's formula (Dix, 1956). The reflection present at a two-way travel time of $t = 65$ ns shows for both broadside and end-fire configuration a well-defined root-mean-square velocity of about $v = 0.13$ m ns^{-1}.

Figure 10(e) and **10(g)** shows TE and TM data measured at the same location after rainfall that look completely different than shown in **Figure 10(a)** and **10(c)**. The reflector at $t = 65$ ns can now only be identified for a limited offset range and is cut off by steeply dipping events enclosed by the black dotted lines. This is due to the increased water content and corresponding increased relative permittivity (see **Table 1**) that results in a total reflection beyond the critical angle θ_{12}^{c} within the surface layer as shown in **Figure 7**. In this way, only trapped waveguide energy can be measured beyond a certain offset. **Figure 10(f)** and **10(h)** shows the corresponding semblance spectrum where the same reflector present at an increased two-way travel time of $t = 70$ ns shows a moderately defined root-mean-square velocity of about $v = 0.12$ m ns^{-1}. The limited offset range where the reflection can be identified results in a reduced resolution of the semblance spectrum. The increased water content caused by the precipitation event clearly results in a reduced average velocity of material above the common reflector, which in turn increases the two-way travel time. Although a maximum in the semblance spectrum indicates the presence of a low-velocity shallow layer in the semblance images in **Figure 10(f)** and **10(h)**, additional semblance analyses using different ranges of offsets lead to different two-way travel time and velocity estimates for this shallow reflection. These variations in travel time and velocity are caused by the presence of interfering multiples and indicate that semblance analysis does not provide reliable estimates of shallow layer properties. Standard velocity analysis of the common reflector at two-way travel time of $t = 75$ ns only reliably provides an average velocity of the medium present above the reflector.

After normalizing the broadside and end-fire data shown in **Figure 10(e)** and **10(g)** on the maximum of each trace, **Figure 11(a)** and **10(b)** shows that most of the energy is present in these steeply dipping events. Due to the geometric spreading being reduced from R^{-1} as typical body waves have (see eqn [11]) to $R^{-1/2}$ since the waves spread their energy in two dimensions, the guided waves can propagate over relatively large distances. The dashed black arrows indicate the apparent phase velocities of these shingling events, whereas the dashed yellow arrows in **Figure 11(a)** and **10(b)** indicate the group velocities with which the energy is traveling through the waveguide. Note that there is a clear difference between the phase and group velocity that indicates a frequency-dependent phase velocity. These properties are characteristic for the presence of dispersive GPR signals in CMP soundings (van der Kruk et al., 2009a). It is clear that these steeply dipping shingling events cannot be fitted with a conventional reflection-type hyperbola.

To determine the frequency-dependent phase velocity, the phase-velocity spectrum can be calculated by transforming the data to the frequency domain and correcting an offset-dependent phase shift for specified ranges of frequency and phase velocity (Park et al., 1999; van der Kruk, 2006) as follows:

$$D(v, f_i) = \left| \sum_{x} \frac{\hat{E}(x, f)}{\left|\hat{E}(x, f)\right|} \exp\left(j\frac{2\pi f_i}{v}x\right) \right| \qquad [31]$$

When a certain frequency is traveling with a certain phase velocity, the formula mentioned earlier results in constructive

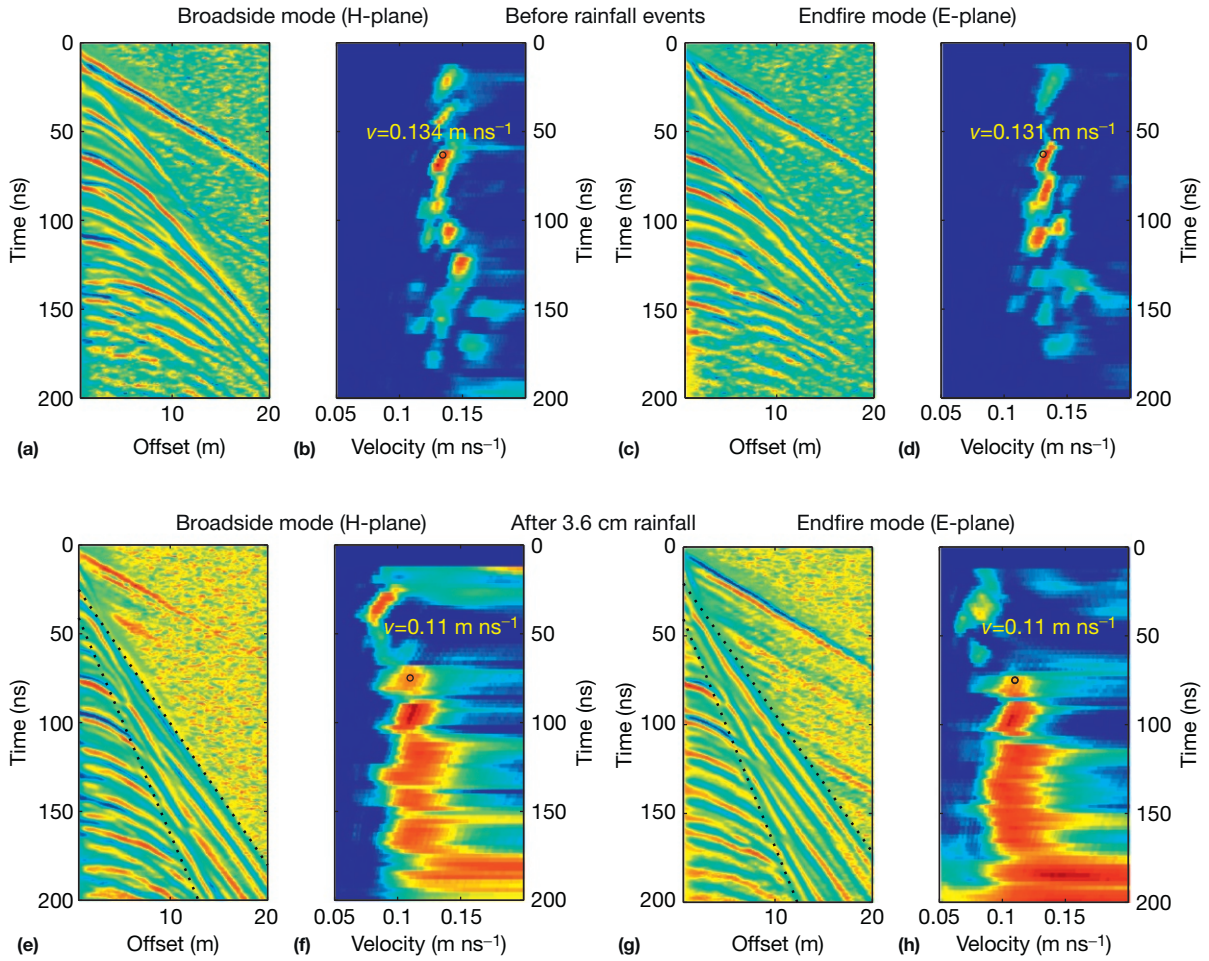


Figure 10 (a) TE and (c) TM common-midpoint profiles measured over a dry surface layer before rainfall and (b) and (d) corresponding velocity semblance spectra. (e) TE and (g) TM CMP profiles measured after 7.2 cm rainfall over a dry surface layer and (f) and (h) corresponding velocity semblance spectra. Picked velocity in the velocity semblance spectra is clearly getting smaller with increasing cumulative water content, which is present in a thin shallow layer. Clear shingling events are present between the dashed lines in (e) and (g) that indicate dispersion due to a shallow low-velocity waveguide. Adapted from van der Kruk J, Jacob RW, and Vereecken H (2010b) Properties of precipitation-induced multilayer surface waveguides derived from inversion of dispersive TE and TM GPR data. *Geophysics* 75: WA263–WA273.

interference of these frequency components that will result in a large spectral amplitude in $D(v,f_i)$, which can be picked to obtain a dispersion curve, $v_{\text{data}}(f_i)=\max[D(v,f_i)]$, that describes the frequency-dependent phase velocity.

The phase-velocity spectra (Park et al., 1999) of the TE and TM data shown in **Figure 11(a)** and **11(b)** are depicted in **Figure 11(c)** and **11(d)**, respectively. Here, three TE and three TM modes can be clearly identified. It is clear from the data shown in **Figures 10** and **11** that the broadside and end-fire data are different because the main contribution for the broadside and end-fire configurations is influenced by the TE and TM reflection/transmission coefficients, respectively, as given in eqns [14]–[15] and [18]–[19]. It can also be seen from **Figure 11(c)–11(f)** that the TE cutoff frequencies (smallest frequency in a mode) are lower than the TM cutoff frequencies as indicated by eqns [26] and [27].

The commonly used seismic techniques to invert dispersive Rayleigh and Love waves (Park et al., 1999; Socco and Strobbia, 2004; Xia et al., 1999, 2004) cannot be used since they are not based on the EM TE and TM reflection coefficients (see eqns [14] and [15]). Recently, techniques for inverting dispersive GPR data have been developed that are based on the fundamental equation of modal theory for EM waves (eqn [25]) and the pertaining EM reflection coefficients.

The medium properties for n layers overlying a half-space, $\varepsilon_1 \ldots \varepsilon_{n+1}$ and $h_1 \ldots h_n$, can be obtained by minimizing the cost function (Strobbia and Cassiani, 2007; van der Kruk, 2006; van der Kruk et al., 2006):

$$C^{\{\text{TE, TM}\}}(\varepsilon_1 \ldots \varepsilon_{n+1}, h_1 \ldots h_n) = \sum_{i=1}^{m} \frac{\left| v_{\text{data}}^{\{\text{TE, TM}\}}(f_i) - v^{\{\text{TE, TM}\}}(f_i, \varepsilon_1 \ldots \varepsilon_{n+1}, h_1 \ldots h_n) \right|}{m} \quad (32)$$

where $v^{\{\text{TE,TM}\}}$ can be calculated using eqn [25]. The cost function associated with multimodal combined TE–TM inversion can be written as (van der Kruk et al., 2010b)

$$C^{\{\text{TE}_0..n\text{TM}_0..m\}} = C^{\text{TE}_0} + \cdots + C^{\text{TE}_n} + C^{\text{TM}_0} + \cdots + C^{\text{TM}_n} \quad [33]$$

The subsurface properties can be inverted using a combined global and local minimization procedure where several starting models are used and different local minima are obtained.

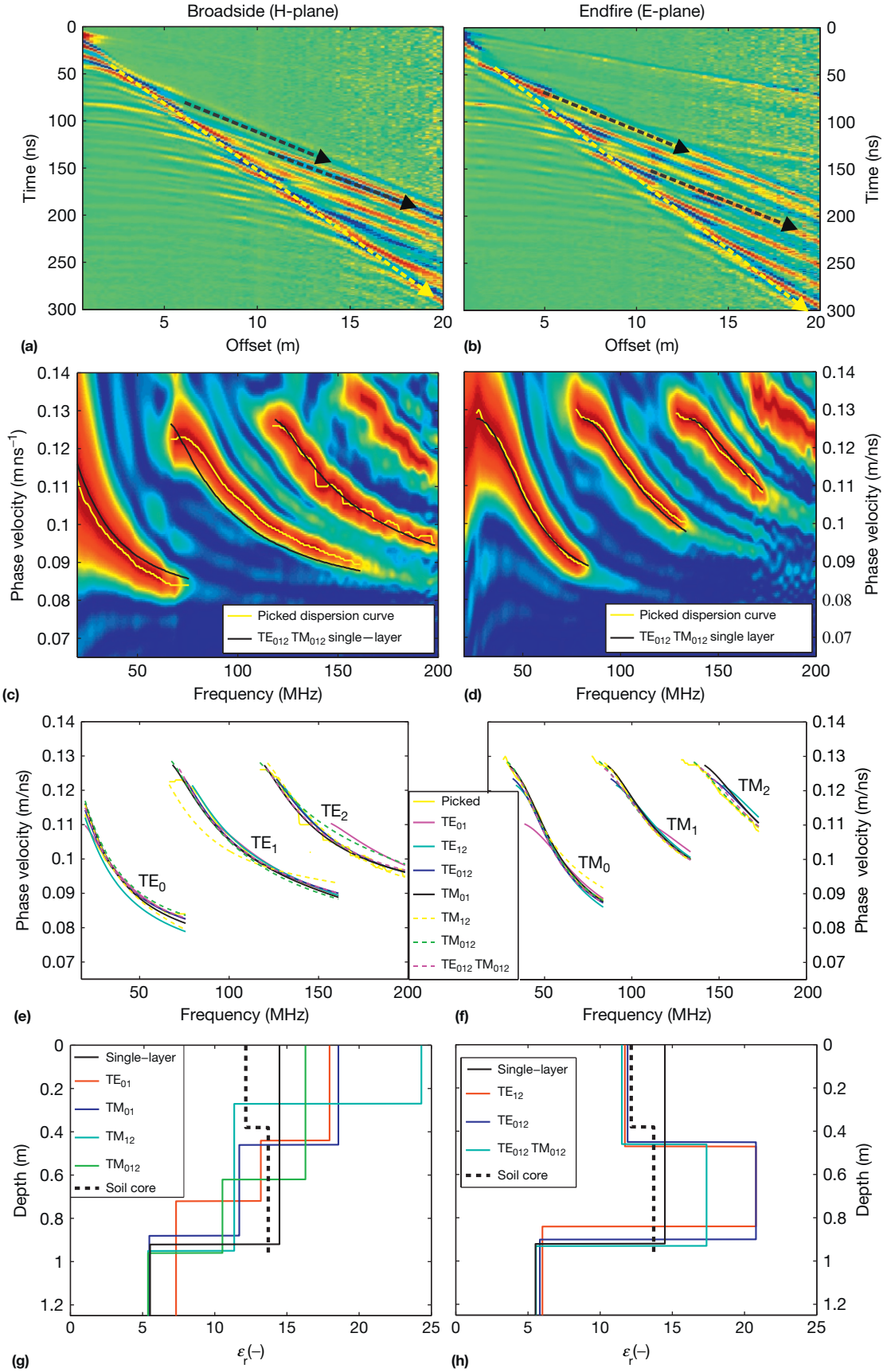


Figure 11 Trace-normalized (a) TE and (b) TM common-midpoint profiles (shown in **Figure 10(c)** and **10(d)**) that show clearly that the dispersive waves (enclosed by the dashed lines) contain most of the energy. The yellow and black dashed arrows are the group and phase velocities, respectively. (c, d) The corresponding TE and TM phase-velocity spectrum and the single-layer $TE_{012}TM_{012}$ inversion results. (e, f) The picked and inverted multilayer dispersion curves and (g, h) show the obtained inverted models and the soil core results (adapted from van der Kruk J, Jacob RW, and Vereecken H (2010b) Properties of precipitation-induced multilayer surface waveguides derived from inversion of dispersive TE and TM GPR data. *Geophysics* 75: WA263–WA273.

The local minimum with best fit to the observed dispersion curve is assumed to be the global minimum, which represents true medium properties (van der Kruk et al., 2006).

Assuming a single-layer waveguide and inverting the data using single and multiple modes return similar medium properties, where the $TE_{012}TM_{012}$ inversion returns $\varepsilon_1 = 14.48$, $\varepsilon_2 = 5.52$, and $h_1 = 0.92$ m (van der Kruk et al., 2010b). The computed TE and TM dispersion curves for this single-layer model are plotted in **Figure 11(c)** and **11(d)** and show that the dispersion curves that are equidistant between the different modes (see also the cutoff frequencies in eqns [26] and [27]) are not able to explain the observed lower-order TE modes.

Since many modes can be identified, a multimode two-layer waveguide inversion using eqns [32] and [33] was carried out that resulted in a better fit compared with the single-layer inversion. For comparison purposes, the computed dispersion curves for the models determined from the multimode inversions are plotted in **Figure 11(e)** and **11(f)**. The multimode inversion results for a two-layer waveguide inversion are shown in **Figure 11(g)** and **11(h)**, where two possible solutions are plotted: a uniformly decreasing permittivity with depth and a high-permittivity middle layer (ε_2), respectively. The TE_{12} and TE_{012} inversions, and the inversion of all data ($TE_{012}TM_{012}$), support the latter interpretation (i.e., a high-permittivity middle layer). A soil core collected at the geophysical test site just prior to collecting the GPR data indicates that there was indeed a high-permittivity middle layer (van der Kruk et al., 2010b). The derived single-layer waveguide height h_1 is close to the sum of the two-layer waveguide heights h_1 and h_2 (see **Figure 11(g)** and **11(h)**). The weighted arithmetic mean of the permittivities ε_1 and ε_2 of the two-layer waveguide, $\varepsilon = 14.07$, is very similar to the single-layer waveguide permittivity $\varepsilon_1 = 14.48$. The single-layer half-space permittivity ε_2 is about the same as the two-layer half-space permittivity ε_3 (van der Kruk et al., 2010). Note that the waveguides in **Figures 8** and **11** have a thickness of about half and twice the wavelength.

More and more data sets that could not be analyzed using ray-based approaches are being identified as containing dispersive waves due to the presence of low-velocity waveguides: a layer of wet, organic silty to gravelly soil overlying a drier layer of sand and gravel (Liu and Arcone, 2003; Arcone et al., 2003), an organic-rich sandy silt overlying gravel units (van der Kruk, 2006; van der Kruk et al., 2006), a mountain slope with a 1 m soil cover (Strobbia and Cassiani, 2007), and thawing of a frozen soil layer (van der Kruk et al., 2009b). Similar inversion techniques have been developed that return thickness and permittivity of the leaky waveguide (van der Kruk et al., 2007a,b). An ice layer overlying water (Arcone, 1984; van der Kruk et al., 2007a,b) and the freezing of wet soil (Steelman et al., 2010; van der Kruk et al., 2009a) are among the CMP data sets that were identified as being strongly dispersive due to leaky waveguides.

11.07.4.3 Common-Offset Processing

Common-offset measurements are mostly carried out for surface and off-ground GPR surveys (see **Figure 1**). In this way, large surveys can be carried out fast and efficiently due to recent developments of accurate positioning systems, like laser theodolite (Heincke et al., 2005; Lehmann and Green, 1999), laser positioning sensors (Grasmueck and Viggiano, 2007), differential GPS (Streich et al., 2006), and self-tracking total station methods (Boniger and Tronicke, 2010). Nowadays, every GPR system is equipped with the possibility to store GPS positioning data directly for every recorded trace such that high-quality and high-precision GPR data can be kinematically acquired using off-the-shelf instrumentation without further hardware modifications.

In several cases, especially when horizontal layers need to be detected in a relatively homogeneous space, it is sufficient to convert the travel time to depth using the obtained wave speed. For dipping layers or small objects, the measured reflections are not originating from the lateral position where they are measured. The process of relocating the reflections to their reflection position is called migration.

For surface GPR surveys, the observed data are commonly processed in a migration-style sense using the Born approximation to obtain the structure and medium properties of the subsurface. Assuming an unbounded homogeneous medium with conductivity σ and permittivity ε that includes a scattering domain of bounded interest with conductivity σ^s and permittivity ε^s, the contrast function χ is given by

$$\chi(x^c) = (\sigma^s - \sigma) + j\omega(\varepsilon^s - \varepsilon) \qquad (34)$$

The electric field that is emitted by a horizontal dipole source oriented in the x_α direction, scattered by the medium contrast function χ, and received by a horizontal dipole receiver oriented in the x_β direction can be described by (Leone and Soldovieri, 2003; Meincke, 2001; van der Kruk et al., 2003a; see also **Figure 9**)

$$\begin{aligned} E^{\text{scat}}_{\alpha\beta}(x^{s,r}, x^c, \omega) &= W(\omega) \int_{V(x^c)} G_{\alpha l}(x^r - x^c, \omega)\chi(x^c)G_{l\beta}(x^c - x^s, \omega)dV \\ &= W(\omega) \int_{V(x^c)} A^{\text{scat}}_{\alpha\beta}(x^{r,s} - x^c, \omega)\exp(-2\gamma R)\chi(x^c)dV \end{aligned} \qquad [35]$$

where the Born approximation is used that is valid for small contrast scatters and $W(\omega)$ is the source wavelet. In this way, a linear expression is obtained that enables a qualitative reconstruction. The first row of the scattering matrix $\hat{E}^{\text{scat}}_{\alpha\beta}$ contains two orthogonal responses of the subsurface due to x_1-oriented source, whereas the second row contains the two orthogonal responses due to an x_2-oriented source. Note that it is also assumed that each point in the subsurface acts as an independent point diffractor and no multiples or mutual interactions between the targets are present in the data.

The directional sensitivity of different antenna configurations show that the far-field response coming from a point scatterer has a maximum in the H-plane and a minimum in the E-plane at the critical angle (e.g., Lehmann et al., 2000; Radzevicius and Daniels, 2000; Roberts and Daniels, 1996; van der Kruk et al., 2003a). Cross polarized antennae do not respond to scatterers located in the vertical planes containing the antenna axes. Instead, the response increases as the scatterers move away from the antenna axes, reaching maxima at azimuths of $\pm 45°$. These results are quite general, because arbitrarily shaped objects can be approximated by ensembles of point scatterers. Using this concept, it can be demonstrated that cross polarized antennae are best able to resolve cylindrical objects when they are oriented at oblique angles to the

antenna axes (Radzevicius and Daniels, 2000). Similarly, it is possible to show that cross polarized antennae will only detect planar boundaries if they strike at an oblique angle to the antenna axes (Streich et al., 2006). Although cross polarized data generally have much lower amplitudes than copolarized data, they may contain significant information (Guy et al., 1999; Orlando and Slob, 2009; Streich et al., 2006).

11.07.4.3.1 Polarization information extraction

By recording the reflected field amplitudes using the four different configurations in the scattering matrix (see eqn [35]), information can be obtained about the orientation of buried objects that have polarization-dependent reflectivity. Polarization effects as functions of the material properties and antenna-to-pipe orientations have been investigated by, for example, Radzevicius and Daniels (2000). Their results show that for metallic pipes and dielectric pipes with permittivities larger than that of the surrounding medium, the direction of preferential scattering usually is aligned with pipe orientation. They also showed that for dielectric pipes with permittivities lower than that of the surrounding medium, the direction of preferential scattering depends on pipe diameter (with respect to the dominant GPR wavelength) and typically is perpendicular to pipe orientation for small pipe diameter-to-wavelength ratios.

After collecting data with four different configurations at a specific orientation to fill the scattering matrix, the Alford rotation (Alford, 1986) enables the reconstruction of the response for every angle of orientation using the following equation:

$$\overline{E}^{\text{scat}}_{\alpha\beta} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \cdot E^{\text{scat}}_{\alpha\beta} \cdot \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \quad [36]$$

The Alford rotation was originally applied to seismic data to determine the main axis of anisotropy and the average azimuthal anisotropy of the medium (e.g., fractures). This approach can also be used to investigate polarization-dependent GPR reflection amplitudes.

Van Gestel and Stoffa (2001) and Seol et al. (2001) had used the Alford rotations of 2-D multicomponent GPR to determine the strike direction of fractures. Sassen and Everett (2009) developed a polarimetric coherency algorithm that is based on the eigenvalue decomposition of the scattering matrix and applied on multicomponent GPR data recorded over a karstic limestone exhibiting a high degree of fracturing. **Figure 12** shows that the multicomponent eigenvalue coherency algorithm in (a) produces the best image of discontinuities in the subsurface compared to pseudoscalar GPR data (Lehmann et al., 2000) in (b) and conventional single-component recording (c). Böniger and Tronicke (2012) introduced an attribute-based analysis approach to extract and characterize buried utility pipes using two broadside antenna configurations based on a symmetry-enhancing image-processing algorithm and principal component analysis to detect and distinguish different utilities present within 3-D GPR data. A depolarization attribute was introduced where a value of 1 indicates a feature exhibiting maximum polarization with the long axes of the antenna dipoles parallel to the long axis of the feature, whereas a value of −1 indicates a feature exhibiting maximum polarization with the long axes of the antenna dipoles perpendicular to the long axis of the feature. **Figure 13** shows a three-dimensional image of pipes with positive depolarization attribute for pipe 2 and negative depolarization attributes for pipes 1, 3, and 4. Comparison with ground-truth data shows that pipes 1, 3, and 4 were from metal and that pipe 2 was from a sewer pipe made of stoneware. The obtained results were consistent with the findings by Radzevicius and Daniels (2000).

11.07.4.3.2 Imaging/migration

The measured reflections are not always coming from below the source and the receiver, but can have traveled from a sphere described by a certain travel time. Especially for a point scatterer, dipping plane, syncline, or anticline, a migration or imaging needs to be carried out to reposition the measured reflections on their actual reflection location in the subsurface. Because of similarities between acoustic and EM processing methods, seismic imaging techniques were initially used for the imaging of GPR data. Several approaches have been implemented in different domains: reverse time migration (e.g., Fisher et al., 1992a; Leuschen and Plumb, 2001), F–K migration (e.g., Bitri and Grandjean, 1998; Fisher et al., 1992b; Grasmueck et al., 1996, 2005), Kirchhoff migration (e.g., Fullagar et al., 2000; Moran et al., 2000; Hammon et al., 2002), synthetic aperture radar (Gilmore et al., 2006), and prestack migration (Feng and Sato, 2004; Leparoux et al., 2001; Pipan et al., 1999).

The amplitudes and phases of raw GPR data depend on the antenna radiation patterns, the vector nature of (EM wave propagation, and the EM properties of the subsurface. Ideally, a GPR imaging algorithm should supply quantitative amplitude information that is related directly to the contrasts in subsurface physical properties by explicitly accounting for the radiation patterns and vector wave propagation. Although several true-amplitude migration techniques exist for seismic data (Gray, 1997), there are few analogous approaches for quantitative imaging of GPR data. Instead, many GPR data sets are migrated using scalar seismic algorithms or remote sensing approaches like synthetic aperture radar (SAR) (Curlander and McDonough, 1991). Such algorithms yield image amplitudes that depend on the antenna radiation patterns and therefore on the relative orientations of the antenna and target reflectors.

Early attempts at quantitative GPR imaging include the schemes of Molyneux and Witten (1993) and Witten et al. (1994), which address these issues for 2-D objects embedded in a homogeneous medium. Other attempts to derive 3-D migration algorithms that include the effects of the antenna radiation patterns and vector wavefields are based on far-field approximations of the radiation patterns. For example, Moran et al. (2000) presented a Kirchhoff migration technique that empirically incorporates the far-field radiation patterns shown in **Figure 5**, and Wang and Oristaglio (2000) used the same patterns in an algorithm based on the generalized Radon transform. Diffraction tomographic approaches, in which the approximations of Green's functions are expressed in the horizontal wave number–frequency domain, have been introduced by Hansen and Johansen (2000) and by Meincke (2001). A linear inverse scattering algorithm for realistic GPR applications was implemented using a singular value decomposition approach (Soldovieri et al., 2007; Soldovieri and

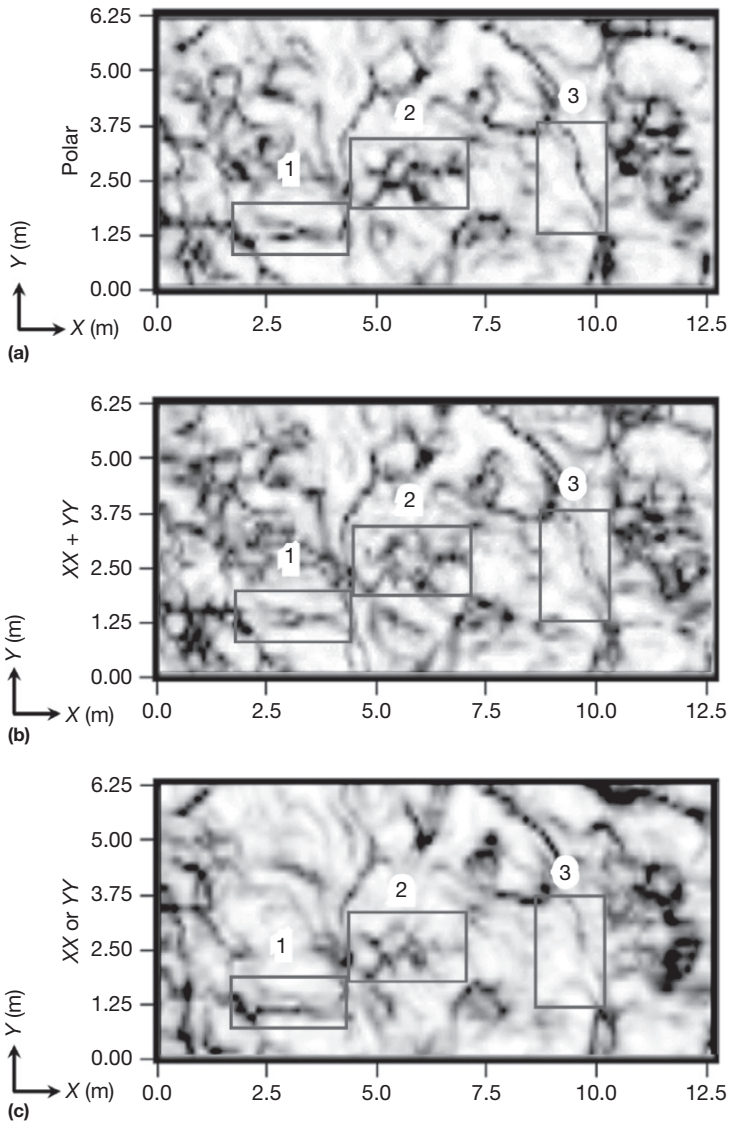


Figure 12 Time slice through a coherency cube created using (a) the full multicomponent data, (b) the 'pseudoscalar' GPR response, and (c) the single component E^{11} (XX) data. Adapted from Sassen DS and Everett ME (2009) 3D polarimetric GPR coherency attributes and full-waveform inversion of transmission data for characterizing fractured rock. *Geophysics* 74: J23–J34.

Solimene, 2007). In the following, the developments to obtain true-amplitude migrated images from GPR data are described.

For a single-component migration, the imaged contrast at a specific depth can be written in frequency domain as

$$\chi_{\alpha\beta}^{\text{image}}(x) = \frac{1}{2\pi}\int \frac{\mathrm{d}\omega}{W(\omega)} \int_{A(x^{r,s})} \hat{H}_{\alpha\beta}(x - x^{r,s}, \omega)\hat{E}_{\alpha\beta}(x^{r,s}, \omega)\mathrm{d}A, \quad [37]$$

where $\hat{H}_{\alpha\beta}$ is the inverse wavefield extrapolator that backpropagates the reflections to their correct subsurface position. The spectral division by $W(\omega)$ is the deconvolution of the source wavelet. The imaging principle of Claerbout (1971) states that the data at zero travel time of the inverse-extrapolated recordings relate to a band-limited version of the physical-property contrasts of the medium of investigation, which is carried out by a summation over ω.

The SAR inverse wavefield extrapolator is defined in the $x-\omega$ domain by the complex conjugate of the phase shift in eqn [6]. Its $k_1-k_2-x_3-\omega$ domain equivalent can be determined using the method of stationary phase (see **Table 2**). The inverse wavefield extrapolator that forms the basis for Gazdag phase-shift migration (see eqn [45] of Gazdag, 1978) is defined in the $k_1-k_2-x_3-\omega$ domain by the phase term $\exp(jk_3^*|x_3|)$, where

$$\begin{aligned} k_3^* &= \sqrt{4k^2 - k_1^2 - k_2^2}, \text{ for } k_1^2 + k_2^2 \le 4k^2 \\ k_3^* &= j\sqrt{k_1^2 + k_2^2 - 4k^2}, \text{ for } k_1^2 + k_2^2 \ge 4k^2 \end{aligned} \quad (38)$$

Ideally, inverse wavefield extrapolation should compensate for all effects of forward propagation described by eqn [35]. This description of the scattered electric field as a convolution of the forward wavefield extrapolator with the contrast

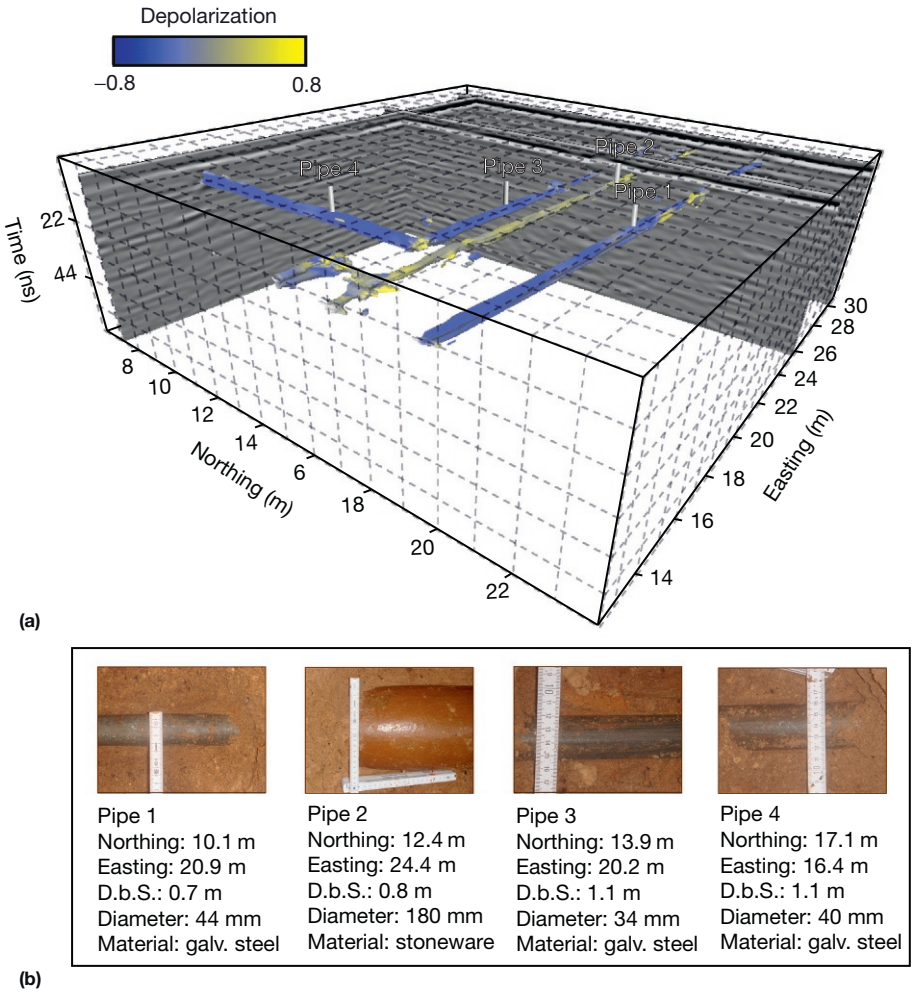


Figure 13 (a) Interpretation of 3-D GPR data where the depolarization attribute is mapped onto the extracted isosurfaces. (b) Ground-truth information obtained from excavation. The white markers in (a) correspond to the pipe locations from the available ground-truth information. Adapted from Böniger U and Tronicke J (2010) On the potential of kinematic GPR surveying using a self-tracking total station: Evaluating system crosstalk and latency. *IEEE Transactions on Geoscience and Remote Sensing* 48: 3792–3798.

Table 2 SAR and Gazdag inverse wavefield extrapolators in the $\mathbf{x}$–ω domain and k_1–k_2–x_3–ω domain

Inverse extrapolator	$\mathbf{x}-\omega$ *domain* $\hat{H}$	$k_1-k_2-x_3-\omega$ *domain* $\dot{H}$
SAR	$\exp(j2kR)$	$\frac{4\pi jk\lvert x_3\rvert}{(k_3^*)^2}\exp(jk_3^*\lvert x_3\rvert)$
Gazdag	$\frac{-jk\lvert x_3\rvert}{\pi R^2}\exp(j2kR)$	$\exp(jk_3^*\lvert x_3\rvert)$

function is converted into a multiplication upon transformation to the $k_1-k_2-x_3-\omega$ domain. Accordingly, in principle, one can readily obtain an inverse operator in the $k_1-k_2-x_3-\omega$ domain by transforming the forward wavefield extrapolator given in eqn [35] to the $k_1-k_2-x_3-\omega$ domain and then taking its inverse. To ensure stability, the inverse wavefield extrapolator is determined only in the propagating wave region; evanescent waves are ignored.

Calculating the inverse wavefield extrapolators in the horizontal frequency domain using the far-field expressions for the electric field in eqn [11] results for individual source–receiver configurations in unbounded operators such that they cannot be used for migration (van der Kruk, 2001). To obtain a suitable approximate inverse of the forward wavefield, the wavefield extrapolators can be combined for the four source–receiver configurations as given in eqn [35], and a multicomponent forward wavefield extrapolator can be introduced that takes into account the vectorial radiation patterns of the dipole antenna and the propagating wavefield (van der Kruk et al., 2003a). Using the exact-field expressions, in which sharp amplitude minima are absent, the multicomponent approach and the formation of inverse wavefield extrapolators for individual source–receiver configurations should result in stable images. However, standard integration routines used to numerically evaluate the integral equations to obtain exact radiation patterns for imaging or migration have long been considered computationally impractical (Radzevicius et al., 2003; Valle et al., 2001; van der Kruk et al., 2003a, b). Recently,

a fast approximate evaluation was introduced for the exact-field half-space expressions at a certain depth level for all horizontal positions simultaneously using an oversampled inverse fast Fourier transformation (Streich and van der Kruk, 2007b). The antenna offset can be accounted for by simple space domain shifting of Green's functions if the offset is a multiple of the spatial sampling and the two-way forward wavefield extrapolator is formed by multiplying the appropriate source and receiver Green's functions. This forward operator now contains the exact radiation patterns of the source and receiver antennae and all propagation effects of the vectorial electric field between the antennae and the assumed reflection point. Each element of the synthetic scattering matrix is then transformed to the $k_1 - k_2 - x_3 - \omega$ domain to allow the inverse wavefield extrapolator to be computed. This can be done by inverting the multicomponent scattering matrix (Streich and van der Kruk, 2007b). Alternatively, with the ability of including exact-field radiation patterns in the 3-D imaging, it is also possible to use a single-component inverse wavefield extrapolator to image single-component data (Streich et al., 2007).

Synthetic tests for a dipping layer show that the amplitudes of the measured E_{11} and E_{22} data depend strongly on the radiation patterns. The amplitudes of the images obtained using the Gazdag approach or far-field radiation patterns for perpendicular source–receiver setups χ_{11} and χ_{22} are different, indicating that the radiation patterns still influence the obtained results process (Streich and van der Kruk, 2007b; Streich et al., 2007, van der Kruk et al., 2010a). The images χ_{11} and χ_{22} using the exact-field radiation patterns are almost invariant to the relative orientations of antenna and reflectors, which shows that the vectorial radiation patterns are removed during the migration process (Streich and van der Kruk, 2007b; Streich et al., 2007).

The true-amplitude vector migration algorithm using the exact field was tested on GPR data acquired at a gravel pit in northern Switzerland (Streich et al., 2007) where an artificial bank of well-sorted gravel was deposited on a dipping surface of less homogeneous natural deposits (**Figure 14(a)**). Two single-component data sets were recorded using copolarized antennae offset by 1 m. The antennae were oriented

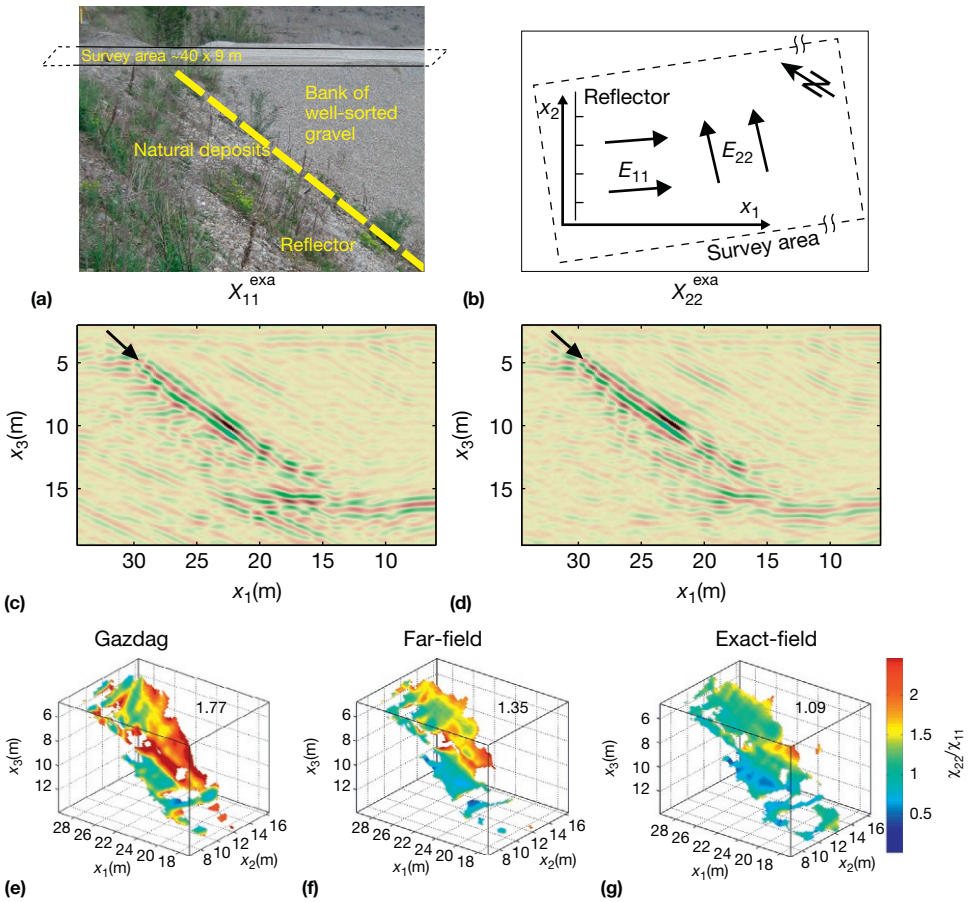


Figure 14 (a) Photograph showing the Weiach survey site (Switzerland). Measurements were made on the planar surface of an artificial gravel bank. The arrow identifies the near-planar dipping boundary between this gravel unit and the underlying natural gravel deposits. (b) Top view, showing the relative orientations of the antennae, survey area (dashed box), and the coordinate system adopted for displaying the results. (c, d) Vertical slices extracted from the 3-D imaged data at $x_2 = 11.25$ m from the χ_{11} and χ_{22} exact-field single-component vector images corresponding to copolarized antennae oriented approximately in the x_1- and x_2-directions, respectively in (b). Similar 3-D data cubes were obtained using the Gazdag and far-field inverse wavefield extrapolator. (e, f, g) The Gazdag, far-field, and exact-field amplitude ratios χ_{22}/χ_{11} for the principal dipping reflector as identified in (c) and (d) by arrows for the exact-field results shown in (g). Modified from Streich R, van der Kruk J, and Green AG (2007) Vector-migration of standard co-polarized 3-D GPR data. *Geophysics* 72: J65–J75.

approximately perpendicular to the reflector trend for the E_{11} data and approximately parallel for the E_{22} data (Figure 14(b)). The reflection from the interface between the artificial bank and the naturally deposited gravels (marked by arrows in Figure 14) is the most prominent event in all images. Whereas the amplitudes of all subsurface reflections differ significantly between the modified scalar χ_{11} and χ_{22} Gazdag images and between the far-field χ_{11} and χ_{22} vector images (Streich et al., 2007), they are similar in the exact-field χ_{11} and χ_{22} vector images (see Figure 14(c) and 14(d)). To assess quantitatively the performance of the different migration schemes, the amplitudes of the principal interface reflection observed in the various χ_{11} and χ_{22} images were compared by picking peak-to-peak amplitudes using the strongest negative troughs (colored brown and marked by arrows in Figure 14) and the following positive peaks in time. The distributions of amplitude ratios χ_{22}/χ_{11} are shown superimposed on the picked reflection surface in Figure 14(e), 14(f), and 14(g). For the modified scalar phase-shift and far-field vector-migrated images, the amplitude analysis yields an average χ_{22}/χ_{11} amplitude ratio of 1.77 and 1.35, respectively. In marked contrast, the amplitude ratio for the exact-field vector-migrated images is reduced to 1.09, which is close to the ideal value of 1. These results indicate that exact-field single-component vector migration has eliminated most radiation pattern and vector wave propagation effects from the field data (Streich et al., 2007).

11.07.4.4 Ray-Based Tomographic Inversion

For crosshole GPR tomography, inversions are generally based on the geometric ray theory. It provides EM velocity and attenuation images of the probed regions by first arrival times and maximum first-cycle amplitude inversions. The domain of interest is divided into cells that have each their own velocity (or relative permittivity) and conductivity. The travel time it takes for the EM waves to travel from a source to a receiver is then described by adding the separate time delays of each cell through which the ray has traveled. The total travel time $\mathbf{t}$ is efficiently described by summing the multiplication of travel paths $\mathbf{L}$ through each cell with the corresponding slowness $\mathbf{u}$ (inverse of velocity) in a matrix form as follows:

$$\begin{bmatrix} \mathbf{t} \\ \lambda\beta\mathbf{u}_0 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{L} \\ \lambda\beta\mathbf{I} \\ \lambda(1-\beta)\mathbf{S} \end{bmatrix} \mathbf{u} \qquad [39]$$

where $\mathbf{S}$ is the smoothing matrix, λ is a smoothing factor, and β is a damping factor, which influence the relative weight of the model regularization with respect to the data and are applied to minimize the amount of deviations from the input model $\mathbf{u}_0$. The slowness $\mathbf{u}$ can be obtained by inverting the former equation using an iterative least squares approach (Holliger et al., 2001) using a conjugate gradient approach. Similar linear equations can be derived for describing the amplitude decay in each cell to estimate conductivity values. Here, far-field expressions are used to correct for the source and receiver radiation patterns (Maurer and Musil, 2004).

Ray-based tomographic schemes are often used for crosshole GPR measurements where many multiple offset gathers are measured to ensure a good ray coverage. Whereas in seismics the tomographic inversion is also often used for surface data, where due to the increasing velocity with depth refracted wave energy can be measured, this approach is less often used for surface GPR data since the velocity is mostly decreasing with depth such that refracted waves are less often present in the data.

Klotzsche et al. (2010) applied the ray-based approaches (eqn [39]) to a crosshole GPR data set acquired at a gravel aquifer in Switzerland (Figure 15(a)). A limited number of source positions were used during the data acquisition and a much larger number of receiver positions were occupied to minimize acquisition time and computational costs. Vertical spacing of 0.5 m and 0.1 m for the source and receiver points were chosen, respectively. Sources and receivers were placed in either borehole to avoid low coverage close to the source borehole. The circles and crosses in Figure 15(c) and 15(d) indicate transmitter and receiver positions. The obtained ray-based permittivity and conductivity (logarithmic scale) images reveal roughly three zones (Figure 15(c) and 15(d)). A zone of higher permittivity and higher electrical conductivity is found in the top 4–5.5 m, followed by lower permittivity and lower conductivity values between 5.5 and 8 m (Klotzsche et al., 2010). The bottom part exhibits intermediate values for both parameters. Within the ray-based inversion results, traces were discarded that contained refracted waves that have traveled through the unsaturated gravel, resulting in a low ray coverage close to the water table (see Figure 15(b)). The measured data and the modeled data using the ray-based inversion results are shown for increasing receiver depths in Figure 15(a) and 15(b), respectively, for sources S1–S11 (see also Figure 14(e)). The measured data for source S3 correspond to the data shown in Figure 3(b). The first arrival times for the synthetic traces that are calculated using the ray-based inversion results are similar, but it is also clear that the modeled data using the ray-based inversion results return a poor match with the measured data and do not fit the phase and amplitudes of the wave trains of the observed data in detail. The limited resolution obtained with this ray-based approach is primarily because only a fraction of the information contained in the radar data (i.e., the first arrival times and maximum first-cycle amplitudes) is included in the inversion.

11.07.4.5 Full-Waveform Inversion

A promising step to improve the inversion of GPR data is to include advanced modeling tools that are able to calculate the propagation of EM waves in complicated medium configurations (Cassidy, 2007; Giannopoulos, 2005; Holliger and Bergmann, 2002; Lampe et al., 2003). Here, the FWI is one of the most promising but also challenging data-fitting techniques (Virieux and Operto, 2009) that is able to use these advanced modeling programs to derive quantitative medium properties. FWI is fairly advanced in exploration seismics (e.g., Mora, 1987; Tarantola, 1984a, 1984b, 1986; Pratt, 1999; Pratt and Ship, 1999; Virieux and Operto, 2009), and recently, three-dimensional inversion algorithms have been implemented (Abubakar et al., 2011; Ali et al., 2008; Vigh and Starr, 2008). Seismic FWI cannot be used to

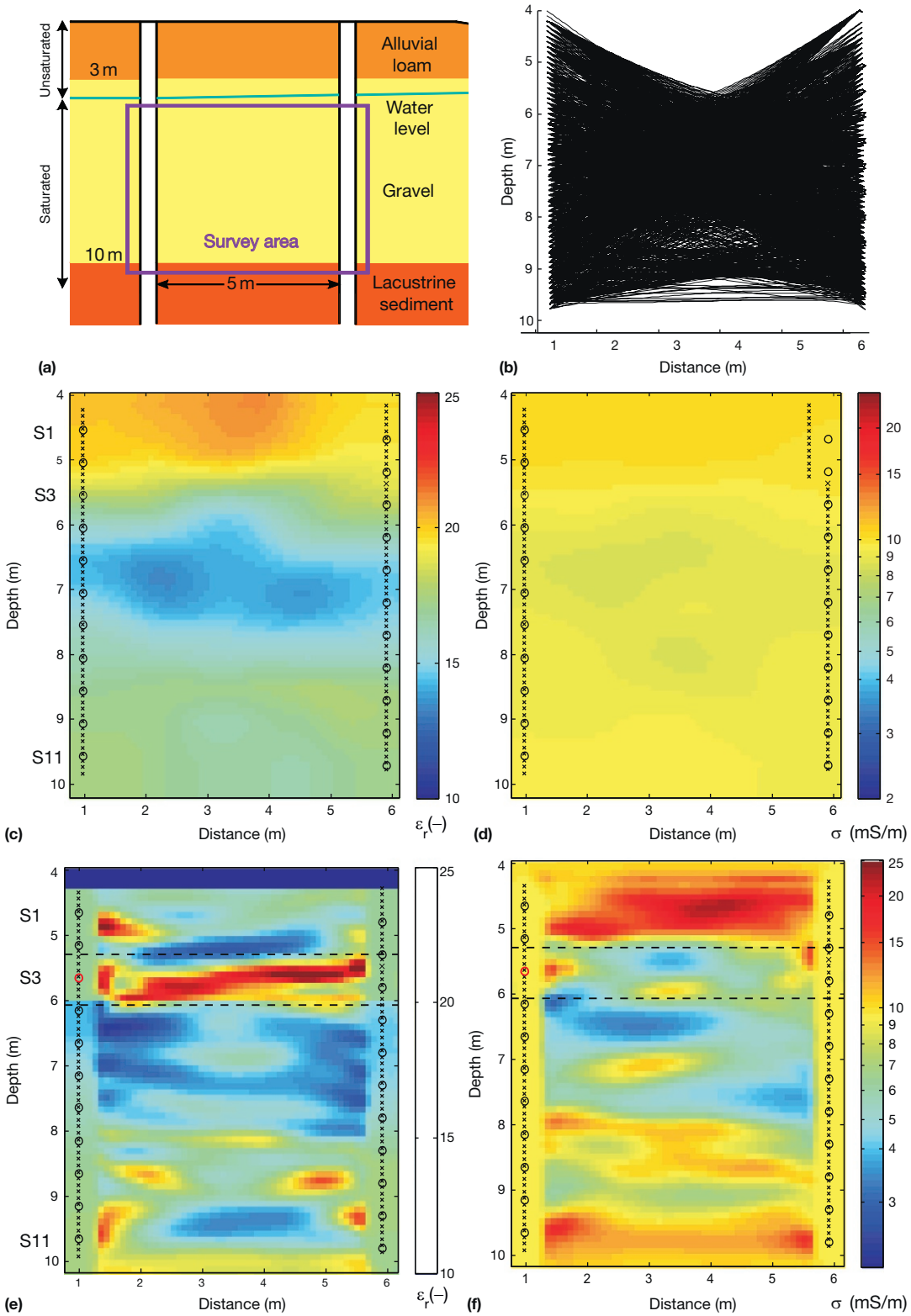


Figure 15 (a) Simplified geologic representation of the experimental area, showing a three-layer structure: alluvial loam, gravel, and lacustrine (clay) sediments (from top to bottom, where the water table is approximately at 4 m depth). (b) Ray coverage based on the travel time inversion shown in (c, d) for the obtained relative permittivity and conductivity distributions. (e, f) The full-waveform permittivity and conductivity inversion results (adapted from Klotzsche A, van der Kruk J, Meles GA, Doetsch J, Maurer H, and Linde N (2010) Full-waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland. *Near Surface Geophysics* 8: 635–649; Klotzsche A, van der Kruk J, Meles GA, and Vereecken H (2012) Crosshole GPR full-waveform inversion of waveguides acting as preferential flow paths within aquifer systems. *Geophysics* 778: H57–H62). The crosses and circles indicate sources and receivers, respectively.

invert GPR data since the pertaining vectorial wave propagation and radiation patterns are different.

Recent developments in FWI of crosshole (Ernst et al., 2007a,b; Kuroda et al., 2007; Klotzsche et al., 2010, 2012, 2013; Meles et al., 2010, 2011), off-ground (Kalogeropoulos et al., 2011, 2013; Lambot et al., 2004), and surface GPR (Busch et al., 2012) use significant information in the measured data and indicate the benefits of FWI inversion approach to estimate quantitative permittivity and conductivity values. A vectorial approach was implemented that honors the vectorial EM wave propagation and simultaneously updates the permittivity and conductivity (Meles et al., 2010). Since in this case the Born approximation (see eqn [35]) is not used, multiples and mutual interactions between the targets are still included in the analysis such that a quantitative analysis of the EM properties is made possible. Nonlinearity issues that especially affect the inverse scattering problems in high-contrast mediums can be avoided by starting the inversion with only the low-frequency content of the data and extending the used bandwidth for later iterations (Meles et al., 2011).

Most full-waveform algorithms use a conjugate gradient method to minimize the cost (misfit) function S in an iterative approach between the modeled $\mathbf{E}^s$ and the measured $\mathbf{E}^s_{\text{obs}}$ traces for all source–receiver combinations within a defined time window. Using the vector formalism introduced by Meles et al. (2010), the misfit function can be written as

$$C(\varepsilon,\sigma)=\frac{1}{2}\sum_s\sum_d\sum_\tau\left[\mathbf{E}^s(\varepsilon,\sigma)-\mathbf{E}^s_{\text{obs}}(\varepsilon,\sigma)\right]^T_{d,\tau}\delta(x-x_d,t-\tau)$$
$$\left[\mathbf{E}^s(\varepsilon,\sigma)-\mathbf{E}^s_{\text{obs}}(\varepsilon,\sigma)\right] \quad [40]$$

where $\mathbf{E}^s$ stands for the electric field of a particular source; T indicates the transposed operator; and s, d, and τ indicate sums over sources, receivers, and observation times. See Meles et al. (2010) for a detailed discussion of this vector formalism. In the first iteration of the FWI, the ray-based inverted ε and σ parameters are used to compute modeled data. The second step estimates the residual wavefield $\mathbf{R}^s$ by subtracting the measured from the modeled data. This residual wavefield $\mathbf{R}^s$ is backpropagated through the model and a zero-lag vectorial cross correlation between it and the forward modeled data is performed to calculate the ε and σ gradients in which the residual is minimized (Meles et al., 2010):

$$\begin{bmatrix}\nabla C_\varepsilon(x')\\ \nabla C_\sigma(x')\end{bmatrix}=\frac{1}{2}\sum_s\frac{(\delta(x-x')\partial_t\mathbf{E}^s)^T\mathbf{G}^T\mathbf{R}^s}{(\delta(x-x')\mathbf{E}^s)^T\mathbf{G}^T\mathbf{R}^s} \quad [41]$$

where $\mathbf{G}^T\mathbf{R}^s$ can be interpreted as the backpropagated residual wavefield. The spatial delta function $\delta(\mathbf{x}-x')$ corresponds to the spatial components of the gradients and reduces the inner product to a zero-lag cross correlation in time. The last step uses appropriate perturbation factors to compute the step lengths needed to update the model parameters

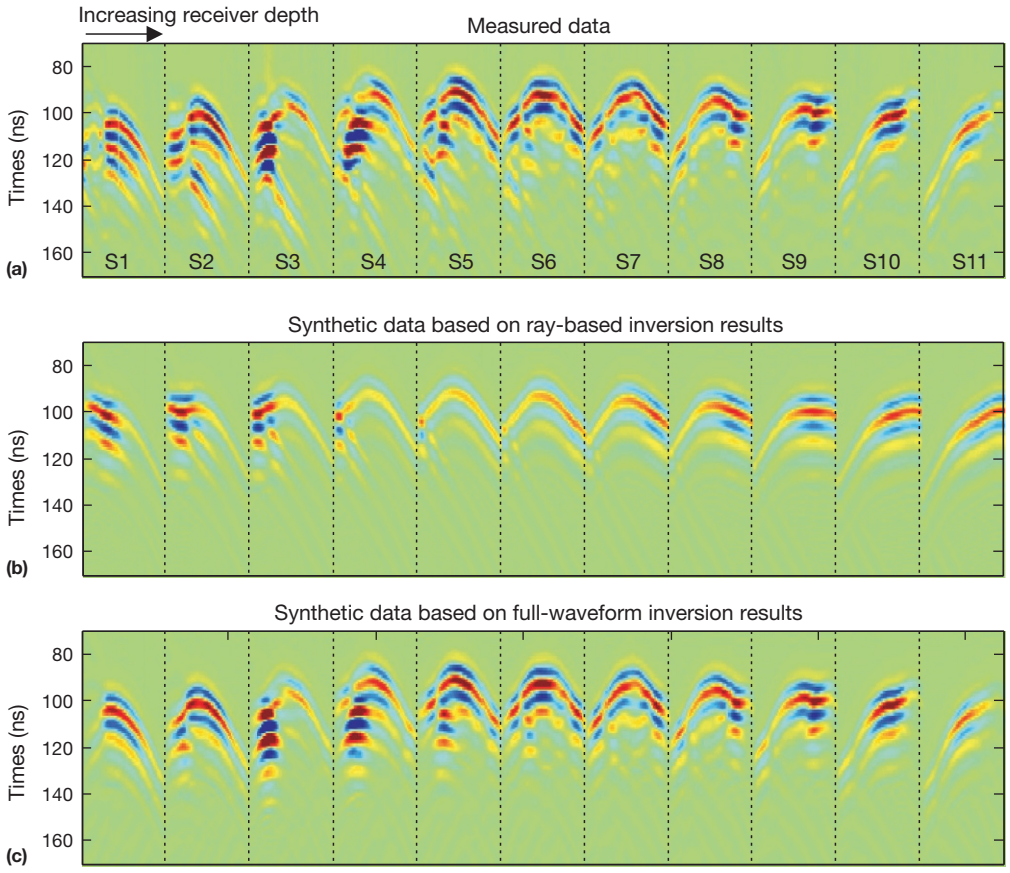


Figure 16 (a) Measured data for representative source–receiver combinations, (b) modeled data using the ray-based inversion results as input, (c) modeled data using the full-waveform inversion (FWI) results as input. Note that all results are showing true amplitude that indicates that the FWI fits very well all details contained in the amplitude and phase of the measured data.

simultaneously. In subsequent iterations, these steps are repeated using the updated model until the residual wavefield is less than a certain percentage (0.5% or 1%) of its value compared to the previous iteration.

For inverting experimental data, a starting model and effective source wavelet needs to be estimated (Klotzsche et al., 2010, 2012) from which a first cut solution is calculated that should match the observed waveforms received in the field experiment within half the dominant pulse period. Otherwise, the inversion will proceed to update in the wrong direction and results in a phenomenon called local minimum trapping (Isernia et al., 2001; Meles et al., 2011). The starting ε and σ tomograms are obtained using standard ray-based inversion (e.g., Holliger et al., 2001; Maurer and Musil, 2004), in which first arrival times and first-cycle amplitudes are used to calculate velocity and attenuation distributions between the boreholes and can then be transformed into ε and σ. Next, the field data are converted from 3-D to 2-D to enable the use of a 2-D FDTD modeling tool. An effective source wavelet that describes the received EM pulse is obtained by first estimating the wavelet shape by using only horizontal rays followed by a deconvolution that returns both wavelet shape and amplitude (Klotzsche et al., 2010).

In the following, an example of FWI is shown for the ray-based inverted crosshole GPR data shown in **Figure 15(a)–15(d)** (Klotzsche et al., 2010, 2012, 2013). Comparing the measured data shown in **Figure 16(a)** for the upper two meters with the modeled data shown in **Figure 16(b)** obtained using the ray-based inversion results (**Figure 15(c)** and **15(d)**) showed an overlap of less than half the dominant pulse period, which is due to the low ray coverage caused by discarding traces that contained refracted waves (see also **Figure 3(b)**). Performing a FWI caused a local minimum trapping and unreliable FWI results for the upper part, which was also indicated by a remaining ε and σ gradient ∇C_ε and ∇C_σ, respectively. Therefore, the ray-based permittivity starting model was adapted by introducing a homogeneous ε_r layer between 4.1 and 6 m in depth and a water table at 4.1 m in the inversion domain. Forward modeling with $\varepsilon_r = 17$ for this upper layer between 4.1 and 6 m in depth showed an overlap of the synthetic radar waveforms with the observed data to within half the dominant pulse period, which is a necessary condition for convergence (Meles et al., 2011).

The final permittivity and conductivity results of the FWI presented in **Figure 15(e)** and **15(f)** show much more details than the ray-based inversion results in **Figure 15(c)** and **15(d)**. The FWI reveals structured detail within these three layers. In particular, the prominent low-velocity layer dipping slightly upward from left to right between the black dashed lines (depth range 5.3–6.1 m) could not be identified in the ray-based inversion results. This thin low-velocity layer is caused by high porosity that was confirmed by flow meter and slug tests (Klotzsche et al., 2013).

In general, the modeled data based on the FWI results shown in **Figure 16(c)** are very similar in shape and amplitude to the measured data (**Figure 16(a)**), which indicates that the obtained models explain the measured data very well. For source S3 that is present within the low-velocity layer (see red circles in **Figure 15(e)** and **15(f)**), the measured data show elongated wave trains and high amplitudes at late arrivals for receivers present at the depth range of the low-velocity zone. The modeled data for source S3 using the inversion results of the FWI show very similar elongated wave trains and high amplitudes at late arrivals as the measured data (compare **Figure 16(c)** with **16(a)**). Note that similar elongated wave trains and high amplitudes can be observed in surface GPR data in the presence of a surface waveguide (see **Figures 8** and **11**).

Table 3 Used symbols

Quantity	*Symbol*	*(Value) SI units*
Levi-Civita tensor	ϵ_{kmp}	–
Vectorial electric field	E_k	V/m
Vectorial magnetic field	H_p	A m
Electric current density	J_k	A/m
Spatial derivative	∂_m	1/m
Temporal derivative	∂_t	1/s
Orthogonal base vectors	$\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{i}_3$	–
Position in three dimensions	$\mathbf{x}$	$x_1\,\boldsymbol{i}_1 + x_2\,\boldsymbol{i}_2 + x_3\,\boldsymbol{i}_3$ m
Permittivity	ε	F/m
Relative permittivity	ε_r	–
Free space permittivity	ε_0	8.854×10^{-12} F/m
Magnetic permeability	μ	H/m
Relative magnetic permeability	μ_r	–
Free space permeability	μ_0	$4\pi\times10^{-7}$ H/m
Electrical conductivity	σ	S m^{-1}
Imaginary unit	j	–
Speed of light in vacuum	c_0	299 792 458 m/s
Electromagnetic wave speed	v	m/s
Green's function	G	1/m
Dyadic Green's function for E as function of J	G^{EJ}_{kr}	A V/m
Propagation constant	γ	1/m
Angular frequency	$\omega = 2\pi f$	rad/s
Frequency	f	Hz
Kronecker delta	δ_{kr}	–
Radius	R	m
Wavelength	λ	m
Polar angle	θ	rad
Azimuthal angle	φ	rad
Wave number	k	1/m
Free space wave impedance	$Z_0 = 120\pi$	Ω
Wave impedance	Z	Ω
Transverse electric reflection coefficient	R^{TE}	–
Transverse magnetic reflection coefficient	R^{TM}	–
Transverse electric transmission coefficient	T^{TE}	–
Transverse magnetic transmission coefficient	T^{TM}	–
(Waveguide) layer thickness	h	m
TE cutoff frequency for the mth mode	$f_c^{\mathrm{TE}_m}$	Hz
TM cutoff frequency for the mth mode	$f_c^{\mathrm{TM}_m}$	Hz
Semblance spectrum	S	–
Source wavelet	W	V
Vertical wave number	k_3	1/m

11.07.5 Conclusions

A multitude of existing and emerging GPR applications show the ability to noninvasively image the subsurface with high

resolution. Due to the rapidly evolving GPR surveying techniques, it is nowadays possible to acquire high-density, high-quality GPR data with high positioning accuracy. Most processing, imaging, and inversion tools are often based on approximations that include far-field, high-frequency ray-based approximations, use a linearized system of equations, assume a two-dimensional space, or use only part of the data that enables a qualitative reconstruction of the subsurface properties.

Several dedicated GPR imaging and inversion techniques explore the TE and TM reflection coefficient variation with offset, the polarization information present in perpendicularly measured broadside data; reconstruct the true-amplitude characteristics of the reflectors; or use the full-waveform information to reconstruct a detailed model of the subsurface. In all these cases, it is crucial to use a detailed forward modeling program that accurately describes the vectorial EM propagation. In this way, these examples are able to go beyond qualitative structural characterization of the subsurface and obtain quantitative information of the subsurface. Due to the increasing computation power, the use of accurate forward modeling programs will become more established such that the full-waveform content of the EM waves can be explored in its full extent for a wide range of studies.

References

Abubakar A, Pan G, Li M, Zhang L, van den Habashy TM, and Berg PM (2011) Three-dimensional seismic full-waveform inversion using the finite-difference contrast source inversion method. *Geophysical Prospecting* 59: 874–888.

Aki K and Richards PG (1980) *Quantitative Seismology*. San Francisco: Freeman & Co.

Alford RM (1986) Shear data in the presence of azimuthal anisotropy. In: *56th Ann. Internat. Mtg., Soc. Expl. Geophys*. pp. 476–479. (expanded abstracts).

Ali HBH, Operto S, and Virieux J (2008) Velocity model building by 3D frequency domain full-.waveform inversion of wide-aperture seismic data. *Geophysics* 73: VE101–VE117.

Annan AP (1973) Radio interferometry depth Sounding1. Theoretical discussion (1973). *Geophysics* 38: 557–580.

Annan AP (2002) GPR—History, trends and future developments. *Subsurface Sensing Technologies and Applications* 3: 253–270.

Annan AP (2005a) Ground-penetrating radar. In: Butler DK (ed.). *Near Surface Geophysics*, pp. 357–438. Tulsa, OK: SEG Press.

Annan AP (2005b) GPR methods for hydrogeological studies. In: Rubin Y and Hubbard S (eds.) *Hydrogeophysics*, pp. 185–214. Dordrecht, Netherlands: Springer.

Arcone SA (1984) Field observations of electromagnetic pulse propagation in dielectric slabs. *Geophysics* 49: 1763–1773.

Arcone SA, Peapples PR, and Liu L (2003) Propagation of a ground-penetrating radar (GPR) pulse in a thin-surface waveguide. *Geophysics* 68: 1922–1933.

Baker GS (1998) Applying AVO analysis to GPR Data. *Geophysical Research Letters* 25: 397–400.

Binley A, Winship P, Middleton R, Pokar M, and West J (2001) High resolution characterization of vadose zone dynamics using cross-borehole radar. *Water Resources Research* 37: 2639–2652.

Bitri A and Grandjean G (1998) Frequency-wavenumber modelling and migration of 2D GPR data in moderately heterogeneous dispersive. *Geophysical Prospecting* 46: 287–301.

Blindow N (2006) Ground penetrating radar. In: Kirsch R (ed.) *Groundwater Geophysics*, pp. 357–438. Tulsa, OK: SEG Press.

Bohidar RN and Hermance JF (2002) The GPR refraction method. *Geophysics* 67: 1474–1485.

Böniger U and Tronicke J (2010) On the potential of kinematic GPR surveying using a self-tracking total station: Evaluating system crosstalk and latency. *IEEE Transactions on Geoscience and Remote Sensing* 48: 3792–3798.

Böniger U and Tronicke J (2012) Subsurface utility extraction and characterization: Combining GPR symmetry and polarization attributes. *IEEE Transactions on Geoscience and Remote Sensing* 50: 736–746.

Bradford JH (2008) Measuring water content heterogeneity using multifold GPR with reflection tomography. *Vadose Zone Journal* 7: 184–193.

Bradford JH and Deeds JC (2006) Ground-penetrating radar theory and application of thin-bed offset-dependent reflectivity. *Geophysics* 71: K47–K57.

Budden KG (1961) *The Wave-Guide Mode Theory of Wave Propagation*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Busch S, van der Kruk J, Bikowski J, and Vereecken H (2012) Quantitative conductivity and permittivity estimation using full-waveform inversion of on-ground GPR data. *Geophysics* 77: H79–H91.

Cassidy NJ (2007) A review of practical numerical modeling methods for the advanced interpretation of ground-penetrating radar in near-surface environments. *Near Surface Geophysics* 5: 5–22.

Claerbout J (1971) Toward a unified theory of reflector mapping. *Geophysics* 36: 467–481.

Cole KS and Cole RS (1941) Dispersion and absorption in dielectrics. *Journal of Chemical Physics* 9: 341–351.

Conyers LB (2011) Discovery, mapping and interpretation of buried cultural resources non-invasively with ground-penetrating radar. *Journal of Geophysics and Engineering* 8: S13–S22.

Conyers LB and Goodman D (1997) *Ground Penetrating Radar: An Introduction for Archaeologists*. Walnut Creek, CA: AltaMira Press, 232 pp.

Curlander J and McDonough R (1991) *Synthetic Aperture Radar, Systems and Signal Processing*. New York: Wiley.

Dai R and Young CT (1997) Transient fields of a horizontal electric dipole on a multilayered dielectric medium. *IEEE Transactions on Antennas and Propagation* 45: 1023–1031.

Daniels DJ (2004) *Ground Penetrating Radar*, 2nd edn. UK: Institution of Engineering and Technology.

Davis JL and Annan AP (1989) Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. *Geophysical Prospecting* 37: 531–551.

de Hoop AT (1995) *Handbook of Radiation and Scattering of Waves*. London: Academic Press Inc.

Deparis J and Garambois S (2009) On the use of dispersive APVO GPR curves for thin-bed properties estimation: Theory and application to fracture characterization. *Geophysics* 74: J1–J12.

Di Matteo A, Pettinelli E, and Slob E (2013) Early-time GPR signal attributes to estimate soil dielectric permittivity: A theoretical study. *IEEE Transactions on Geoscience and Remote Sensing* 51: 1643–1654.

Dix CH (1956) Seismic velocities from surface measurements. *Geophysics* 20: 68–86.

Endres AL, Clement WP, and Rudolph DL (2000) Ground penetrating radar imaging of an aquifer during a pumping test. *Ground Water* 38: 566–576.

Engheta N, Papas CH, and Elachi C (1982) Radiation patterns of interfacial dipole antennas. *Radio Science* 17: 1557–1566.

Ernst JR, Green AG, Maurer H, and Holliger K (2007) Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data. *Geophysics* 72: J53–J64.

Ernst JR, Maurer H, Green AG, and Holliger K (2007) Full-waveform inversion of cross-hole radar data based on 2D finite-difference time domain solutions of Maxwell's equations. *IEEE Transactions on Geoscience and Remote Sensing* 45: 2807–2828.

Feng X and Sato M (2004) Pre-stack migration applied to GPR for landmine detection. *Inverse Problems* 20: S99–S115.

Fisher E, McMechan GA, and Annan AP (1992) Acquisition and processing of wide-aperture ground-penetrating radar data (1992). *Geophysics* 57: 495–504.

Fisher E, McMechan GA, Annan AP, and Cosway SW (1992) Examples of reverse-time migration of single-channel, ground-penetrating radar profiles. *Geophysics* 57: 577–586.

Fullagar PK, Livelybrooks DW, Zhang P, et al. (2000) Radio tomography and borehole radar delineation of the McConnell nickel sulfide deposit, Sudbury, Ontario, Canada. *Geophysics* 65: 1920–1930.

Galagedara LW, Parkin GW, and Redman JD (2003) An analysis of the ground-penetrating radar direct ground wave method for soil water content measurement. *Hydrological Processes* 17: 3615–3628.

Garambois S, Senechal P, and Perroud H (2002) On the use of combined geophysical methods to assess water content and water conductivity of near-surface formations. *Journal of Hydrology* 259: 32–48.

Gazdag J (1978) Wave equation migration with the phase-shift method. *Geophysics* 43: 1342–1351.

Giannopoulos A (2005) Modelling ground penetrating radar by GprMax. *IEEE Construction and Building Materials* 19: 755–762.

Gilmore C, Jeffrey I, and Lovetri J (2006) Derivation and comparison of SAR and frequency-wavenumber migration within a common inverse scalar wave problem formulation. *IEEE Transactions on Geoscience and Remote Sensing* 44: 1454–1460.

Grasmueck M (1996) 3-D ground-penetrating radar applied to fracture imaging in gneiss. *Geophysics* 61: 1050–1064.
Grasmueck M and Viggiano DA (2007) Integration of ground-penetrating radar and laser position sensors for real-time 3-D data fusion. *IEEE Transactions on Geoscience and Remote Sensing* 45: 130–137.
Grasmueck M, Weger R, and Horstmeyer H (2005) Full-resolution 3D GPR imaging. *Geophysics* 70: K12–K19.
Gray S (1997) True-amplitude seismic migration: A comparison of three approaches. *Geophysics* 62: 929–936.
Greaves RJ, Lesmes DP, Lee JM, and Toksoz MN (1996) Velocity variations and water content estimated from multi-offset, ground-penetrating radar. *Geophysics* 61: 683–695.
Gross RA, Green AG, Horstmeyer H, Holliger K, and Baldwin J (2003) 3-D georadar images of an active fault: Efficient data acquisition, processing and interpretation strategies. *Subsurface Sensing Technologies and Applications* 4: 19–40.
Guy ED, Daniels JJ, Radzevicius SJ, and Vendl MA (1999) Demonstration of using crossed dipole GPR antennae for site characterization. *Geophysical Research Letters* 26: 3421–3424.
Hammon WS, Zeng XX, Corbeanu RM, et al. (2002) Estimation of the spatial distribution of fluid permeability from surface and tomographic GPR data and core, with a 2-D example from the Ferron Sandstone, Utah. *Geophysics* 67: 1505–1515.
Hansen T and Johansen P (2000) Inversion scheme for ground penetrating radar that takes into account the planar air–soil interface. *IEEE Transactions on Geoscience and Remote Sensing* 38: 496–506.
Heincke H, Green AG, van der Kruk J, and Horstmeyer H (2005) Acquisition and processing strategies for 3-D georadar surveying a region characterized by rugged topography. *Geophysics* 70: K53–K61.
Hollender F and Tillard S (1998) Modeling ground-penetrating radar wave propagation and reflection with the Jonscher parameterization. *Geophysics* 63: 1933–1942.
Holliger K and Bergmann T (1998) Accurate and efficient modeling of ground-penetrating radar antenna radiation. *Geophysical Research Letters* 25: 3883–3886.
Holliger K and Bergmann T (2002) Numerical modeling of borehole georadar data. *Geophysics* 67: 1249–1257.
Holliger K, Musil M, and Maurer H (2001) Ray-based amplitude tomography for cross-hole georadar data: A numerical assessment. *Journal of Applied Geophysics* 47: 285–298.
Hugenschmidt J (2002) Concrete bridge inspection with a mobile GPR system. *Construction and Building Materials* 16: 147–154.
Huisman JA, Hubbard SS, Redman JD, and Annan AP (2003) Measuring soil water content with ground penetrating radar: A review. *Vadose Zone Journal* 2: 476–491.
Isernia T, Pascazio V, and Pierri R (2001) On the local minima in a tomographic imaging technique. *IEEE Transactions on Geoscience and Remote Sensing* 59: 1596–1607.
Jacob RW and Hermance JF (2004) Assessing the precision of GPR velocity and vertical two-way travel time estimates. *Journal of Environmental and Engineering Geophysics* 9: 143–153.
Jadoon KZ, Weihermueller L, Scharnagl B, et al. (2012) Estimation of soil hydraulic parameters in the field by integrated hydrogeophysical inversion of time-lapse ground-penetrating radar data. *Vadose Zone Journal* 11. http://dx.doi.org/10.2136/vzj2011.0177.
Jol HM (2009) *Ground Penetrating Radar Theory and Applications*. Amsterdam: Elsevier Science.
Jonscher AK (1977) The universal dielectric response. *Nature* 267: 673–679.
Kalogeropoulos A, van der Kruk J, Hugenschmidt J, Bikowski J, and Brühwiler E (2013) Full-waveform GPR inversion to assess chloride gradients in concrete. *NDT&E International* 57: 74–84.
Kalogeropoulos A, van der Kruk J, Hugenschmidt J, Busch S, and Merz K (2011) Chlorides and moisture assessment in concrete by GPR full waveform inversion. *Near Surface Geophysics* 9: 277–285.
Klotzsche A, van der Kruk J, Linde N, Doetsch J, and Vereecken H (2013) 3-D characterization of high-permeability zones in a gravel aquifer using 2-D crosshole GPR full-waveform inversion and waveguide detection. *Geophysical Journal International* 195: 932–944. http://dx.doi.org/10.1093/gji/ggt275.
Klotzsche A, van der Kruk J, Meles GA, Doetsch J, Maurer H, and Linde N (2010) Full-waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland. *Near Surface Geophysics* 8: 635–649.
Klotzsche A, van der Kruk J, Meles GA, and Vereecken H (2012) Crosshole GPR full-waveform inversion of waveguides acting as preferential flow paths within aquifer systems. *Geophysics* 778: H57–H62.
Knight R (2001) Ground penetrating radar for environmental applications. *Annual Review of Earth and Planetary Sciences* 29: 229–255.
Kong JA (1986) *Electromagnetic Wave Theory*. New York: John Wiley.
Kuroda S, Takeuchi M, and Kim HJ (2007) Full-waveform inversion algorithm for interpreting crosshole radar data: A theoretical approach. *Korean Geosciences Journal* 11: 211–217.
Lambot S, Slob EC, van den Bosch E, Stockbroeckx B, Scheers B, and Vanclooster M (2004) Estimating soil electric properties from monostatic ground-penetrating radar signal inversion in the frequency domain. *Water Resources Research* 40, W04205.
Lambot S, Weihermuller L, Huisman JA, et al. (2006) Analysis of air-launched ground-penetrating radar techniques to measure the soil surface water content. *Water Resources Research* 42, W11403.
Lampe B and Holliger K (2005) Resistively loaded antennas for ground penetrating radar: A modeling approach. *Geophysics* 70: K23–K32.
Lampe B, Holliger K, and Green AG (2003) A finite-difference time-domain simulation tool for ground-penetrating radar antennas. *Geophysics* 68: 971–987.
Lehmann F, Boerner D, Holliger K, and Green A (2000) Multicomponent georadar data: Some implications for data acquisition and processing. *Geophysics* 65: 1542–1552.
Lehmann F and Green A (1999) Semi-automated georadar data acquisition in three dimensions. *Geophysics* 64: 836–848.
Leone G and Soldovieri F (2003) Analysis of the distorted born approximation for subsurface reconstruction: Truncation and uncertainties effects. *IEEE Transactions on Geoscience and Remote Sensing* 41: 66–74.
Leparoux D, Gibert D, and Cote P (2001) Adaptation of prestack migration to multi-offset ground-penetrating radar (GPR) data. *Geophysical Prospecting* 49: 374–386.
Leuschen CJ and Plumb RG (2001) A matched-filter-based reverse-time migration algorithm for ground-penetrating radar data. *IEEE Transactions on Geoscience and Remote Sensing* 39: 929–936.
Liu L and Arcone SA (2003) Numerical simulation of the wave-guide effect of the near surface thin layer on radar wave propagation. *Journal of Environmental and Engineering Geophysics* 8: 133–141.
Maurer H and Musil M (2004) Effects and removal of systematic errors in crosshole georadar attenuation tomography. *Journal of Applied Geophysics* 55: 261–270.
Meincke P (2001) Linear GPR inversion for lossy soil and a planar air-soil interface. *IEEE Transactions on Geoscience and Remote Sensing* 39: 2713–2721.
Meles G, Greenhalgh S, van der Kruk J, Green A, and Maurer H (2011) Taming the non-linearity problem in GPR full-waveform inversion for high contrast media. *Journal of Applied Geophysics* 73: 174–186.
Meles G, van der Kruk J, Greenhalgh SA, Ernst JA, Maurer H, and Green AG (2010) A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination cross-hole/borehole-to-surface GPR data. *IEEE Transactions on Geoscience and Remote Sensing* 48: 3391–3407.
Michalski KA and Mosig JR (1997) Multilayered media Green's functions in integral equation formulations. *IEEE Transactions on Antennas and Propagation* 45: 508–519.
Molyneux JE and Witten A (1993) Diffraction tomographic imaging in a monostatic measurement geometry. *IEEE Transactions on Geoscience and Remote Sensing* 31: 507–511.
Mora P (1987) Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics* 52: 1211–1228.
Moran MR, Greenfield S, Arcone S, and Delaney A (2000) Multidimensional GPR array processing using Kirchhoff migration. *Journal of Applied Geophysics* 43: 281–295.
Neal A (2004) Ground—Penetrating radar and its use in sedimentology: principles, problems and progress. *Earth-Science Reviews* 66: 261–330.
Orlando L and Slob E (2009) Using multicomponent GPR to monitor cracks in a historical building. *Journal of Applied Geophysics* 67: 327–334.
Park CB, Miller RD, and Xia J (1999) Multichannel analysis of surface waves. *Geophysics* 64: 800–808.
Peters L, Daniels JJ, and Young JD (1994) Ground-penetrating radar as a subsurface environmental sensing tool. *Proceedings of the IEEE* 82: 1802–1822.
Pipan M, Baradello L, Forte E, et al. (1999) 2-D and 3-D processing and interpretation of multi-fold ground penetrating radar data: A case history from an archaeological site. *Journal of Applied Geophysics* 41: 271–292.
Pratt RG (1999) Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics* 64: 888–901.
Pratt RG and Shipp RM (1999) Seismic waveform inversion in the frequency domain, part 2: Fault delineation in sediments using crosshole data. *Geophysics* 64: 902–914.
Radzevicius SJ, Chen CC, Peters L Jr., and Daniels JJ (2003) Near-field dipole radiation dynamics through FDTD modeling. *Journal of Applied Geophysics* 52: 75–91.
Radzevicius SJ and Daniels JJ (2000) Ground penetrating radar polarization and scattering from cylinders. *Journal of Applied Geophysics* 45: 111–125.

Redman JD, Davis JL, Galagedara LW, et al. (2002) Field studies of GPR air launched surface reflectivity measurements of soil water content. *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE)* 4758: 156–161.
Reppert PM, Morgan FD, and Toksöz MN (2000) Dielectric constant determination using ground-penetrating radar reflection coefficients. *Journal of Applied Geophysics* 43: 189–197.
Roberts RL and Daniels JJ (1996) Analysis of GPR Polarization Phenomena. *Journal of Environmental and Engineering Geophysics* 1: 139–157.
Saarenketo T and Scullion T (2000) Road evaluation with ground penetrating radar. *Journal of Applied Geophysics* 43: 119–138.
Sambridge M (1999a) Geophysical inversion with a neighbourhood algorithm. I. Searching a parameter space. *Geophysical Journal International* 138: 479–494.
Sambridge M (1999b) Geophysical inversion with a neighbourhood algorithm. II. Apprising the ensemble. *Geophysical Journal International* 138: 727–746.
Sassen DS and Everett ME (2009) 3D polarimetric GPR coherency attributes and full-waveform inversion of transmission data for characterizing fractured rock. *Geophysics* 74: J23–J34.
Seol SJ, Kim JK, Song Y, and Chung SH (2001) Finding the strike direction of fractures using GPR. *Geophysical Prospecting* 49: 200–308.
Slob EC and Fokkema JT (2002a) Interfacial dipoles and radiated energy. *Subsurface Sensing Technologies and Applications* 3: 347–367.
Slob EC and Fokkema JT (2002b) Coupling effects of two electric dipoles on an interface. *Radio Science* 37: 1073.
Slob EC, Sato M, and Olhoeft G (2010) Surface and borehole ground-penetrating-radar developments. *Geophysics* 75: A103–A120.
Smith GS (1984) Directive properties of antennas for transmission into a material half-space. *IEEE Transactions on Antennas and Propagation* 32: 232–246.
Socco LV and Strobbia C (2004) Surface-wave method for near-surface characterization: A tutorial. *Near Surface Geophysics* 2: 165–185.
Soldovieri F, Hugenschmidt J, Persico R, and Leone G (2007) A linear inverse scattering algorithm for realistic GPR applications. *Near Surface Geophysics* 5: 14–29.
Soldovieri F and Solimene R (2007) Through-wall imaging via a linear inverse scattering algorithm. *IEEE Transactions on Geoscience and Remote Sensing* 4: 513–517.
Steelman CM and Endres AL (2010) An examination of direct ground wave soil moisture monitoring over an annual cycle of soil conditions. *Water Resources Research* 46, W11533.
Steelman CS, Endres AL, and van der Kruk J (2010) Field observations of shallow freeze and thaw processes using high-frequency ground-penetrating radar. *Hydrological Processes* 24: 2022–2033.
Streich R and van der Kruk J (2007a) Characterizing a GPR antenna system by near-field electric field measurements. *Geophysics* 72: A51–A55.
Streich R and van der Kruk J (2007b) Accurate imaging of multicomponent GPR data based on exact radiation patterns. *IEEE Transactions on Geoscience and Remote Sensing* 45: 93–103.
Streich R, van der Kruk J, and Green AG (2006) Three-dimensional multicomponent georadar imaging of sedimentary structures. *Near Surface Geophysics* 4: 39–48.
Streich R, van der Kruk J, and Green AG (2007) Vector-migration of standard co-polarized 3-D GPR data. *Geophysics* 72: J65–J75.
Strobbia C and Cassiani G (2007) Multilayer ground-penetrating radar guided waves in shallow soil layers for estimating soil water content. *Geophysics* 72: J17–J29.
Taner MT and Koehler F (1969) Velocity Spectra—Digital computer derivation and application of velocity functions. *Geophysics* 34: 859–881.
Tarantola A (1984a) Inversion of seismic-reflection data in the acoustic approximation. *Geophysics* 49: 1259–1266.
Tarantola A (1984b) Linearized inversion of seismic reflection data. *Geophysical Prospecting* 32: 998–1015.
Tarantola A (1986) A strategy for non-linear elastic inversion of seismic-reflection data. *Geophysics* 51: 1893–1903.
Tillard S and Dubois JC (1995) Analysis of GPR data—Wave-propagation velocity determination. *Journal of Applied Geophysics* 33: 77–91.
Tsoflias GP and Becker MW (2008) Ground-penetrating-radar response to fracture-fluid salinity: Why lower frequencies are favorable for resolving salinity changes. *Geophysics* 73: J25–J30.
Tsoflias GP and Hoch A (2006) Investigating multi-polarization GPR wave transmission through thin layers: Implications for vertical fracture characterization. *Geophysical Research Letters* 33, L20401.
Tsoflias GP, Van Gestel JP, Stoffa PL, Blankenship DD, and Sen M (2004) Vertical fracture detection by exploiting the polarization properties of ground-penetrating radar signals. *Geophysics* 69: 803–810.
Valle S, Zanzi L, and Rocca F (1999) Radar tomography for NDT: Comparison of techniques. *Journal of Applied Geophysics* 41: 259–269.
van der Kruk J (2001) Three-dimensional imaging of multi-component ground penetrating radar. PhD Dissertation, Delft University of Technology.
van der Kruk J (2004) Three-dimensional GPR imaging in the horizontal wavenumber domain for different heights of source and receiver antennas. *Near Surface Geophysics* 2: 25–32.
van der Kruk J, Streich R, and Grasmueck M (2010a) Towards true-amplitude vector migration of GPR data using exact radiation patterns. In: Miller RD, Bradford JD, and Holliger K (eds.) *Advances in Near Surface Seismology and Ground Penetrating Radar Book*. *Geophysical Developments Series 15,* pp. 97–116. Tulsa, OK: Society of Exploration Geophysicists.
van der Kruk J, Jacob RW, and Vereecken H (2010b) Properties of precipitation-induced multilayer surface waveguides derived from inversion of dispersive TE and TM GPR data. *Geophysics* 75: WA263–WA273.
van der Kruk J (2006) Properties of surface waveguides derived from inversion of fundamental and higher mode dispersive GPR data. *IEEE Transactions on Geoscience and Remote Sensing* 44: 2908–2915.
van der Kruk J, Arcone SA, and Liu L (2007) Fundamental and higher mode inversion of dispersed GPR waves propagating in an ice layer. *IEEE Transactions on Geoscience and Remote Sensing* 45: 2483–2491.
van der Kruk J, Jacob RW, and Vereecken H (2009a) Identifying dispersive GPR signals and inverting for surface wave-guide properties. *The Leading Edge* 28: 1234–1239.
van der Kruk J, Steelman CM, Endres AL, and Vereecken H (2009b) Dispersion inversion of electromagnetic pulse propagation within freezing and thawing soil waveguides. *Geophysical Research Letters* 36: L18503–L18507.
van der Kruk J, Streich R, and Green AG (2006) Properties of surface waveguides derived from separate and joint inversion of dispersive TE and TM GPR data. *Geophysics* 71: K19–K29.
van der Kruk J, Wapenaar CPA, Fokkema JT, and van den Berg PM (2003a) Three-dimensional imaging of multicomponent ground-penetrating radar data. *Geophysics* 68: 1241–1254.
van der Kruk J, Wapenaar CPA, Fokkema JT, and van den Berg PM (2003b) Improved three-dimensional image reconstruction technique for multicomponent ground-penetrating radar data. *Subsurface Sensing Technologies and Applications* 4: 61–99.
van Gestel JP and Stoffa PL (2001) Application of Alford rotation to ground-penetrating radar data. *Geophysics* 66: 1781–1792.
van Overmeeren RA, Sariowan SV, and Gehrels JC (1997) Ground penetrating radar for determining volumetric soil water content; Results of comparative measurements at two test sites. *Journal of Hydrology* 197: 316–338.
Vigh D and Starr EW (2008) 3D prestack plane-wave, full-waveform inversion. *Geophysics* 73: 135–144.
Virieux J and Operto S (2009) An overview of full-waveform inversion in exploration geophysics. *Geophysics* 74: WCC1–WCC26.
Wang TL and Oristaglio ML (2000) GPR imaging using the generalized Radon transform. *Geophysics* 65: 1553–1559.
Weihermüller L, Huisman JA, Lambot S, Herbst M, and Vereecken H (2007) Mapping the spatial variation of soil water content at the field scale with different ground penetrating radar techniques. *Journal of Hydrology* 340: 205–216.
Witten AJ, Molyneux JE, and Nyquist JE (1994) Ground penetrating radar tomography: Algorithms and case studies. *IEEE Transactions on Geoscience and Remote Sensing* 32: 461–467.
Wu TT and King RP (1965) The cylindrical antenna with non-reflecting resistive loading. *IEEE Transactions on Antennas and Propagation* 13: 369–373.
Xia J, Miller RD, and Park CB (1999) Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics* 64: 691–700.
Xia J, Miller RD, Park CB, and Ivanov J (2004) Utilization of high-frequency Rayleigh waves in near-surface geophysics. *The Leading Edge* 23: 753–759.
Yilmaz O (2000) *Seismic Data Analysis—Processing, Inversion, and Interpretation of Seismic Data*. Tulsa, OK: SEG.
Zeng X, McMechan GA, and Xu T (2000) Synthesis of amplitude-versus-offset variations in ground penetrating radar data. *Geophysics* 65: 113–125.

11.08 Tools and Techniques: Electrical Methods

A Binley, Lancaster University, Lancaster, UK

© 2015 Elsevier B.V. All rights reserved.

11.08.1	**Introduction**	233
11.08.2	**Measurement Principles**	234
11.08.2.1	DC Resistivity	234
11.08.2.1.1	Current flow in a homogenous earth	234
11.08.2.1.2	In situ measurement of resistivity from the ground surface	234
11.08.2.1.3	Measurement practicalities	235
11.08.2.1.4	Electrode array geometry	235
11.08.2.1.5	Sensitivity to lateral variability in resistivity	236
11.08.2.1.6	Sensitivity to azimuthal variability in resistivity	237
11.08.2.1.7	Resistivity measurements in layered media	237
11.08.2.2	Induced Polarization	238
11.08.2.2.1	Chargeability	239
11.08.2.2.2	Complex resistivity/conductivity	240
11.08.2.3	Measurement Errors	242
11.08.3	**Field Configuration**	242
11.08.3.1	1-D Electrical Sounding	243
11.08.3.2	2-D Surface-Based Imaging	243
11.08.3.3	3-D Surface-Based Imaging	245
11.08.3.4	Borehole Logging	245
11.08.3.5	Borehole-Based Imaging	245
11.08.3.6	Optimal Measurement Configuration	247
11.08.3.7	Monitoring and Automated Data Collection	247
11.08.4	**Modeling and Inversion**	248
11.08.4.1	Forward Modeling	248
11.08.4.1.1	IP forward modeling	249
11.08.4.2	Inverse Modeling	249
11.08.4.2.1	3-D inversion	250
11.08.4.2.2	Time-lapse inversion	251
11.08.4.2.3	IP inversion	251
11.08.4.2.4	Accounting for other a priori information in an inversion	251
11.08.4.2.5	Image appraisal	253
11.08.4.2.6	Data and model errors	254
11.08.5	**Summary and Outlook**	255
Acknowledgments		256
References		256

11.08.1 Introduction

In this chapter, we present a range of field investigative techniques for the determination of subsurface electrical properties. Here, the term electrical properties is used to define low-frequency (<1 kHz) conductive and capacitive characteristics. Electrical methods are sometimes referred to as geoelectrical methods; here, we focus on methods utilizing an active source (self-potential, i.e., passive source; methods are described in Chapter X). In the majority of investigations, direct current (DC) resistivity (or its inverse, DC conductivity) is the focus of study. Although less common, methods for the determination of conductive and capacitive properties are deployed for a range of applications, in some cases for the purpose of assessing electrical spectroscopic characteristics (i.e., the variation in properties as a function of the frequency of injected current).

Modern electrical methods are widely assumed to originate from the work in 1912 by Conrad Schlumberger, who developed a means of metallic ore prospecting using a four-electrode array. However, as Vernon (2008) noted, British pioneers Alfred Williams and Leo Daft developed an electrical prospecting approach in the late nineteenth century and in 1901 formed 'The Electrical Ore-Finding Company Ltd,' which unfortunately proved to have limited success, despite attempts to apply their approach internationally. The company went into receivership in 1905; however, the Daft and Williams equipment and method were adopted and developed in Sweden (Dahlin, 2001) and elsewhere. Schlumberger, however, truly pioneered a successful electrical prospecting method, and throughout the first half of the twentieth century, electrical methods remained relatively similar to the original Schlumberger approach. In the latter part of the twentieth century,

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00192-5

instrument developments permitting multiplexed, automatically switched, electrode arrays, coupled with the development of robust inversion tools and widespread availability of appropriate computational resources, led to a rapid growth in the use of electrical imaging systems, albeit still based on measurements similar to the original Schlumberger concepts. During this period, electrical imaging approaches, very similar to the geophysical methods, were advancing in fields of process (chemical) engineering (York, 2001) and biomedical imaging (Barber and Brown, 1984; Wexler et al., 1985).

Electrical methods are probably the most widely used near-surface geophysical techniques, certainly for environmental investigations. This is because (1) subsurface electrical properties are often well correlated to physical and chemical properties of fluids within the pore space (e.g., saturation and salinity) and lithologic properties (e.g., porosity and clay content) (Glover, Chapter 11.04); (2) the theoretical concepts are relatively straightforward; (3) field measurement techniques are highly scalable, allowing investigations to depths of tens of centimeters to hundreds of meters; (4) instrumentation is relatively low cost and, for DC resistivity at least, straightforward to operate; and (5) data analysis techniques have matured – robust data inversion tools, for example, are widely available.

In this chapter, we describe the principles behind the range of techniques available, illustrate the measurement configurations, and outline the concepts behind data analysis and inversion tools.

11.08.2 Measurement Principles

11.08.2.1 DC Resistivity

11.08.2.1.1 Current flow in a homogenous earth

In Chapter 11.04, we defined resistivity using the example of uniform electrical current flow through a sample. In a field environment, since we are not able to create a uniform pathway of current, we need an alternative way of deducing the resistivity. The general approach involves injecting current between two electrodes acting as a point source/sink and then measuring the voltage drop between two other electrodes.

For a three-dimensional (3-D), isotropic, electrical resistivity distribution, $\rho(\mathbf{x})$, the electric potential, $V(\mathbf{x})$, due to a single (point) current electrode, with strength I, is defined by

$$\nabla \cdot \left(\frac{1}{\rho}\nabla V\right) = -I\delta(\mathbf{x}) \qquad [1]$$

where $\nabla = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}$ and δ is the Dirac delta function.

Equation [1] is subject to boundary conditions

$$\left(\frac{1}{\rho}\right)\frac{\partial V}{\partial n} + \beta V = 0 \qquad [2]$$

where n is the outward normal and β defines the type of boundary condition ($\beta = 0$: Neumann; β nonzero: mixed). Equation [1] thus allows the definition of boundary conditions in terms of a flux, a voltage, or a combination of the two.

Let us first consider a subsurface of uniform resistivity and the case where current is injected between two electrodes on the ground surface, one of which is so far away from the region of interest that it acts as an infinite electrode. Assuming a current sink at infinity, the solution of eqn [1] for this example results in an expression for the voltage at some distance r from the current source:

$$V = \frac{I\rho}{2\pi r} \qquad [3]$$

11.08.2.1.2 In situ measurement of resistivity from the ground surface

Taking the concept in the preceding text, we can now consider a practical means of determining resistivity in the field. We focus initially on measurements conducted on the ground surface but show later how these principles can be translated to more general configurations, for example, within, or between, boreholes.

If we move the current sink electrode to the region of investigation, the resulting potential field is more focused. If we then measure the voltage difference between two points, then we can compute the resistivity of the subsurface. The current electrode configuration is referred to as a dipole and the potential electrode pair (the 'receiver') is often referred to as a measurement dipole. The term 'quadrapole' is commonly used as a label for the four-electrode combination. The labels A, B, M, and N are usually used as labels for the two current and two potential electrodes, respectively, as shown in Figure 1.

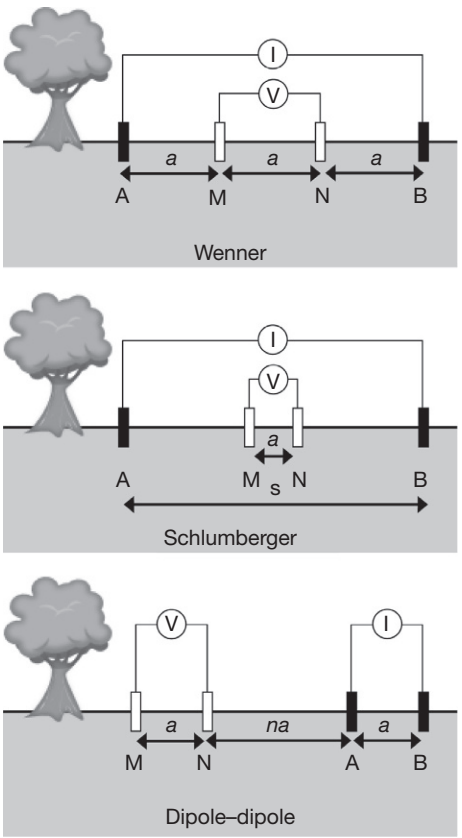


Figure 1 Wenner, Schlumberger, and, dipole–dipole configurations.

From the method of superposition, using eqn [3], we can write an expression for the voltage at any point as

$$V = \frac{I\rho}{2\pi r_{C^+}} - \frac{I\rho}{2\pi r_{C^-}} \quad [4]$$

where r_{C^+} is the distance from the C^+ electrode and r_{C^-} is the distance from the C^- electrode. Therefore, if we consider the potential dipole, we can compute the voltage difference as

$$\Delta V = \frac{I\rho}{2\pi}\left(\frac{1}{AM} - \frac{1}{BM} - \frac{1}{AN} + \frac{1}{BN}\right) \quad [5]$$

where AM is the distance between A and M electrodes, AN is the distance between A and N electrodes, etc.

We can write eqn [5] in terms of resistivity:

$$\rho_a = \frac{k\Delta V}{I} \quad [6]$$

where k is a term called the geometric factor, which can be written as

$$k = \frac{2\pi}{\left(\frac{1}{AM} - \frac{1}{BM} - \frac{1}{AN} + \frac{1}{BN}\right)} \quad [7]$$

Notice that in eqn [6], we have used ρ_a as the symbol for resistivity. ρ_a is referred to as the apparent resistivity to indicate that this is the value that is apparent from the measurement. For a homogenous, flat earth, this value will be the true resistivity of the subsurface. The apparent resistivity is a convenient measure because it has the same units as resistivity and allows us to assess measurements independently of the current injected in the ground. However, as described later, inverse methods must be applied to apparent resistivity data in order to determine the resistivity structure where variability exists.

11.08.2.1.3 Measurement practicalities

In order to measure the apparent resistivity, we need a current source (electrodes A and B in Figure 1) and a way of measuring the voltage between electrodes M and N. For ground-based measurements, the electrodes are typically stainless steel stakes. The size of the stakes will depend on the survey scale (remember that we have assumed that they act as point sources and sensors). For most near-surface studies, electrodes that are typically 1 cm in diameter, 30 cm long, and pushed in the ground no more than 10 cm are adequate. The use of stainless steel prevents corrosion on the electrodes, which can be problematic for the potential electrodes as chemical reactions will lead to electrochemical signals that could influence the measurements. In addition, any coating on the electrode will influence the contact between the electrode and the soil. The electrical resistance across this contact should not be too high (say less than a few kΩ); otherwise, the voltage we measure will be influenced. To avoid polarization of the potential electrodes, nonpolarizing electrodes may be used. This is often not necessary for DC resistivity measurements, but is important for induced polarization (IP) surveys, as discussed later. Modern resistivity instruments attempt to combat the polarization of the electrodes by regularly switching the polarity of the current electrodes A and B; that is, an alternating current is used. This is often in the form of a switched square wave with a typical time period of around 1 s or more.

Since we are interested in the voltage difference relative to the injected current, the resistivity measurement is, therefore, really a resistance measurement, and the instruments are often referred to as earth resistance meters.

11.08.2.1.4 Electrode array geometry

The geometric configuration of the electrode quadrapole shown in the preceding text and the derivation of an apparent resistivity are quite generalized. There are, however, specific geometries that are widely used in practice. As shown in Figure 1, in the Wenner array configuration, the distances AM, MN, and BN are all equal, to some value that we will call a. If we apply eqn [7] in order to get the geometric factor k, we can see that the apparent resistivity for this configuration can be expressed as

$$\rho_a = \frac{2\pi a\Delta V}{I} \quad [8]$$

In the Schlumberger array configuration (see Figure 1), the current electrodes A and B still remain outside the potential electrodes M and N (thus keeping the measured potential difference high), but the distance between the electrodes M and N is much smaller than that between the current electrodes (distance MN $< 0.2 \times$ distance AB). In this case and again using the general expression for the geometric factor (eqn [7]), the apparent resistivity can be written as

$$\rho_a = \frac{\pi(s^2 - a^2)}{4a}\frac{\Delta V}{I} \quad [9]$$

where s is the distance AB and a is the distance MN.

In the dipole–dipole array configuration, the current electrode dipole is adjacent to the potential electrode dipole, and the dipoles are equal in width, a, and separated by distance na (as shown in Figure 1). The separation na is used to indicate an integer (n) multiplier of the dipole width. In this case, the apparent resistivity can be written as

$$\rho_a = \pi n(n+1)(n+2)a\frac{\Delta V}{I} \quad [10]$$

Note that in this configuration, the measured voltage gradient will be reduced in comparison with the Wenner configuration, for example, with the same electrode spacing.

Although the Wenner, Schlumberger, and dipole–dipole arrays are the most common configurations for ground-based studies, a wide range of other electrode arrangements have been utilized. In the pole–dipole array, the current 'sink' (electrode B) is placed at some remote location (making it equivalent to an infinite electrode), giving a geometric factor of $2\pi((AM \times AN)/(AN - AM))$. In the square array, current (and potential) electrodes are adjacent to each other, resulting in a geometric factor of $\pi a(2+\sqrt{2})$, where a is the size of the square. Such an array is potentially valuable for assessing anisotropy of resistivity (e.g., Habberjam, 1972).

The various geometries have advantages and disadvantages, and the choice of configuration should be based on the intended application and expected signal strength. Figure 2 shows example sensitivity patterns for the Wenner and dipole–dipole configurations. In these examples, sensitivity is computed as

$$\text{Sensitivity} = \frac{\partial \log(\rho_a)}{\partial \log(\rho)} \quad [11]$$

This gives a measure of how the observed apparent resistivity is sensitive to changes in variation of resistivity of the subsurface.

These patterns reveal the zone of influence (depth and lateral extent) for a given measurement. The two examples shown in Figure 2 highlight the potential high lateral sensitivity of the dipole–dipole array, in contrast to the Wenner configuration. Note that sensitivities may be positive or negative; that is, an increase in observed apparent resistivity may be due to increases in resistivity in part of the region or decreases in other parts. The effect of this is demonstrated in examples showing the sensitivity to lateral variability in resistivity.

Studies of DC resistivity sensitivity patterns are reported in Roy and Apparao (1971), Barker (1979), and Blome (2010). Roy and Apparao (1971) provide a depth of investigation for various four-electrode arrays; their analysis suggests that suitable depths for Wenner, Schlumberger, and dipole–dipole arrays are (in increasing order of vertical resolution) $0.11L$, $0.125L$, and $0.18L$, respectively, where L is the longest distance between electrodes. It should be noted that these depths are based on an assumption of uniform resistivity, which will clearly not be the case in a real situation.

The signal strength is also an important consideration when selecting measurement arrays: the dipole–dipole configuration can result in weak voltage gradients at the receiver location, particularly for large separation between transmitter pair and receiver pair; in contrast, the Wenner and Schlumberger arrays provide relatively stronger signals due to the location of the receiver pair within the current electrode pair. High-quality (sensitive) receiver components and high-voltage sources (transmitter) may help to overcome some of the difficulties associated with poor signal-to-noise configurations but can add additional resource and field safety constraints. The Schlumberger array is widely used for depth profiles (soundings), as discussed later. Besides logistic benefits, its advantage over other arrays, for soundings, is the weak sensitivity to lateral variability, which is a considerable advantage when focusing the field investigation to the exploration of vertical variability in resistivity. Table 1 summarizes some important characteristics of three common electrode arrays.

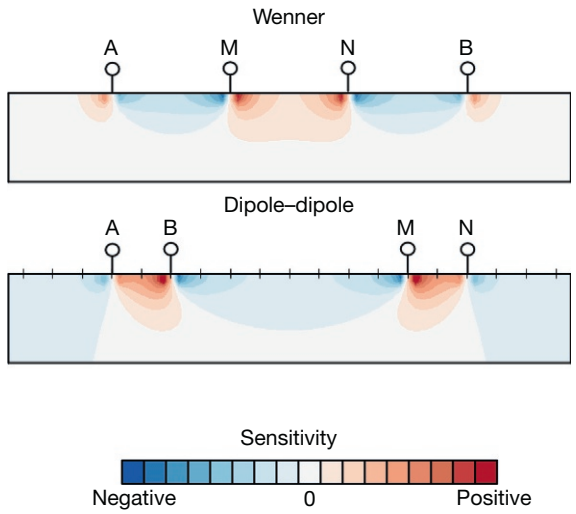


Figure 2 Sensitivity patterns, for example, Wenner and dipole–dipole configurations in a uniform resistivity.

Traditionally, geoelectrical surveys were conducted with a specific array configuration (Wenner, Schlumberger, etc.). As discussed later, the development of computer-controlled electrode switching tools has permitted easy deployment of multiple or hybrid measurement schemes. More recently, attempts have been made to develop measurement configurations optimized for the specific field problem. We discuss these developments later, particularly in the context of 2-D and 3-D imaging, but first, consider some fundamental characteristics of DC resistivity measurements for simple problems.

11.08.2.1.5 Sensitivity to lateral variability in resistivity

We start with a simple case in which the resistivity varies only laterally, as shown in Figure 3. This could represent a fault divide, which demarks a boundary of an aquifer, although we use it here simply to show the sensitivity of the resistivity measurement to the contrast in the subsurface. First, we consider a dipole–dipole array electrode configuration (given its superior lateral sensitivity) with an electrode spacing (a) of 10 m and $n=1$ and then compute the apparent resistivities we would expect to observe as we moved the array across the divide. (Note that these have been computed using a numerical model, since it is not possible to derive them using an analytic expression.)

Figure 3 shows how the apparent resistivity observed will vary using the 10 m-spaced array as it is moved laterally along the surface. Clearly, in the far left and right regions of the survey, the measurements will be insensitive to the resistivity on the other side of the divide, but as the array crosses the divide, we observe a variation in apparent resistivity as the measurement samples both units. Given that the sensitivity field is quite complex (Figure 2, shown for a uniform resistivity), we do not see a linear change in the apparent resistivity. Notice also that even when the entire array is in one of the regions (on the left or right of the divide), we see some sensitivity to the contrasting resistivity of the subsurface. Again, this is a logical consequence of the potential field created by the current dipoles and the resultant sensitivity pattern. We can illustrate this further with a wider-spaced electrode array in the aforementioned example. Figure 3 shows how the apparent resistivities measured with a 20 m-spaced electrode array will vary across the divide. As expected, in

Table 1 Comparison of common electrode arrays for electrical resistivity measurements

	Wenner	*Schlumberger*	*Dipole–dipole*
Depth of investigation	L	M	H
Vertical resolution	H	M	L
Signal strength	H	M	L
Suitability for vertical sounding	M	H	L
Suitability for lateral profiling	M	L	H

Characteristics are shown. H, high; M, medium; L, low.

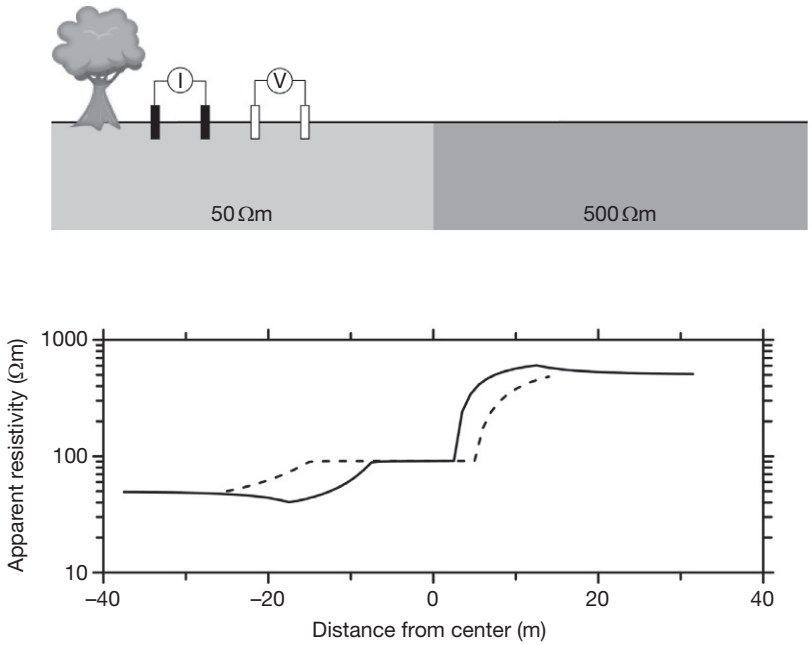


Figure 3 Example for illustration of sensitivity to lateral changes in resistivity. The solid line indicates the measured apparent resistivity for a 10 m-spaced dipole; the dashed line shows the equivalent for a 20 m-spaced dipole.

comparison with the 10 m-spaced array, we see a greater influence of the contrast in resistivity further away from the divide.

The example in the preceding text not only shows the effectiveness of resistivity measurements for mapping lateral variability in subsurface resistivity but also highlights the fact that the measured apparent resistivities will be influenced by the geometry of the electrode array and that, within the region of contrast in resistivity, the values cannot be used to infer the true resistivity of the subsurface. Clearly, in a more realistic system with more complex lateral variations, we will expect to see even more artifacts in the observed data. Nevertheless, this simple approach can be extremely effective for mapping large areas. This use of the method is often called profiling since it is normally used along transects. The approach is extremely popular for archaeological surveys for mapping shallow subsurface features (e.g., Clark, 1990), although in this case, the computation of true resistivity values is often unnecessary as it is the areas of contrast and their structure (e.g., linear bodies representing buried parts of ancient buildings), that is, the main focus of investigation. In fact, the archaeological geophysics community have adapted a different electrode array configuration for this type of study: this is the twin array (e.g., Gaffney and Gater, 2003), which consists of a remote B (current sink) electrode and a remote N (voltage reference) electrode, coupled with roaming A and M electrodes (the 'twin') that are separated by a fixed distance (typically 1 m). This allows much easier maneuverability of the electrode array across the site of interest.

11.08.2.1.6 Sensitivity to azimuthal variability in resistivity

A variant on the profile method is an azimuthal survey in which colinear four-electrode measurements (typically Wenner or Schlumberger) are made at different horizontal angles, centered on a single point. By rotating the electrode array in this way, it is possible to determine any change in the apparent resistivity with orientation. Such surveys can be useful for determining, for example, the orientation of fracture planes (see, e.g., Lane et al., 1995; Taylor and Flemming, 1988; Watson and Barker, 1999). **Figure 4** illustrates the approach. The apparent resistivity will be a function of angle (θ), and such data would normally be presented in a polar plot. Note that as electrode spacing increases in this example, we would expect to see a change in the apparent resistivity values as the fracture planes are not vertical.

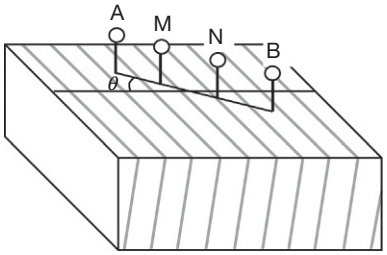


Figure 4 Illustration of azimuthal measurements in anisotropic media. The gray lines represent dipping fracture planes.

11.08.2.1.7 Resistivity measurements in layered media

We now consider resistivity measurements in layered media, which will help to not only illustrate further the effect of sensitivity to contrasts in resistivity but also reveal more about the effect of the distribution of the potential field on the area of measurement sensitivity. Investigations of layered systems in hydrogeology, for example, are extremely common given that many sedimentary hydrologic systems exhibit vertical contrasts in physical properties. We start by illustrating the

effect of a two-layered resistivity structure on the observed apparent resistivity.

11.08.2.1.7.1 Two-layered media

Consider the subsurface structure shown in Figure 5. Here, a 10 m thick low-resistivity (50 Ωm) unit overlies a more resistive (500 Ωm) unit, which extends to an infinite depth. If we compute the current paths for two-electrode spacings for this case, we can see two important effects. First, the current paths for the shorter array spacing are concentrated near the surface, and thus, the measurement of apparent resistivity at this spacing will not be influenced by the deeper resistivity unit. Second, for the longer electrode spacing, the current paths penetrate deeper, as expected, but we see that the near-surface low-resistivity unit in the layered case has an effect of reducing the depth of current flow. These characteristics are important in understanding the sensitivity to the location and characteristic (the resistivity) of layers in the subsurface.

Figure 5(c) shows how the apparent resistivity observed from a Schlumberger electrode array on the surface will vary with current electrode spacing for the case in Figure 5(a) and 5(b) (recall that the potential electrode dipole lies in the center of the current dipole). As the array gets larger, the effect of the near-surface layer is reduced (as the volume of investigation covers a greater proportion of the deeper unit), but the upper layer has an impact on the apparent resistivity, even at very large electrode separations. It is important to recognize that the depth (and volume) that we 'sense' with the measurement is influenced not only by the geometry of the electrode array and the type of array (Wenner, Schlumberger, etc.) but also by the resistivity structure of the subsurface under investigation.

11.08.2.1.7.2 Three-layered media

Let us now consider the case of a three-layered resistivity structure. Figure 6(a) shows a three-layered system comprising of a 5 m thick shallow resistive (500 Ωm) unit, overlying a 10 m thick layer of lower resistivity (200 Ωm), which overlies a deep unit, for which we will assign two contrasting resistivities (50 and 500 Ωm). Comparing the apparent resistivity to electrode spacing for these two cases (Figure 6(b) and 6(c)), we can see how the sensitivity to the middle layer changes with the resistivity of the deeper unit: for the low-resistivity case, the measurements appear insensitive to the middle layer because the current is drawn more into the deeper layer (Figure 6(b)).

The examples so far illustrate how the apparent resistivity varies as a function of lateral and vertical variation in subsurface resistivity. In order to determine an assessment of the true resistivity structure, data processing is necessary. This is addressed later in the chapter, in Section 11.08.4.

11.08.2.2 Induced Polarization

The DC resistivity method provides information about the electrical conductive properties of the subsurface. In contrast, the IP method targets the capacitive characteristics, which can give additional insight into the physical and electrochemical nature of subsurface materials (Glover, Chapter 11.04). The IP method is effectively an extension of the DC resistivity

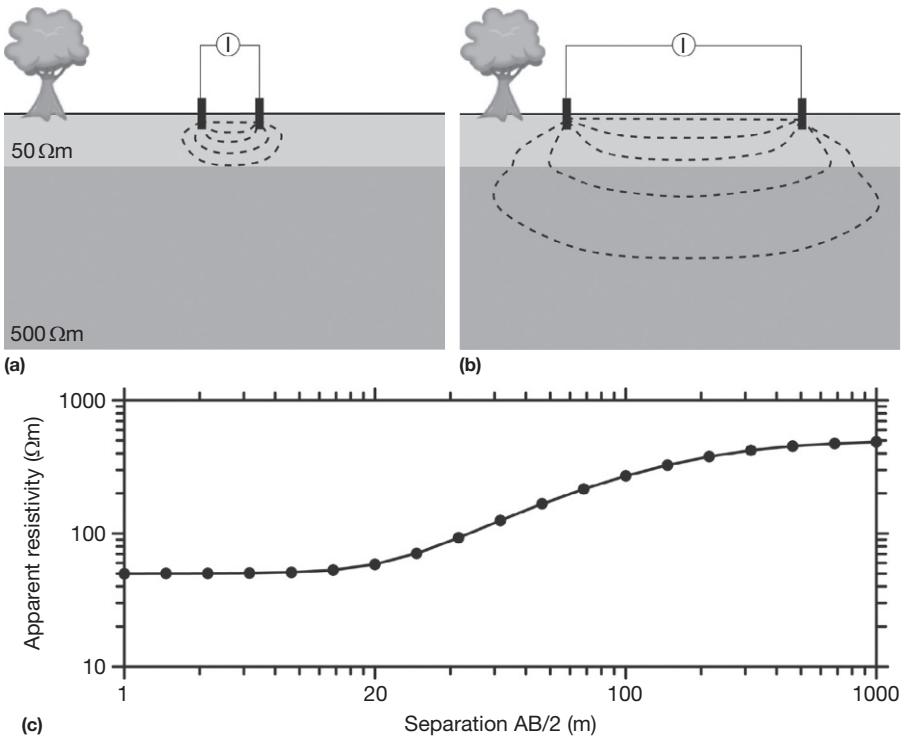


Figure 5 Effect of vertical variation in resistivity on measured response. A 10 m thick 50 Ωm unit overlies an infinitely thick 500 Ωm unit. (a) and (b) show the effect of vertical variation in resistivity on current paths for two different current dipole spacings; (c) shows computed apparent resistivity for a Schlumberger array sounding.

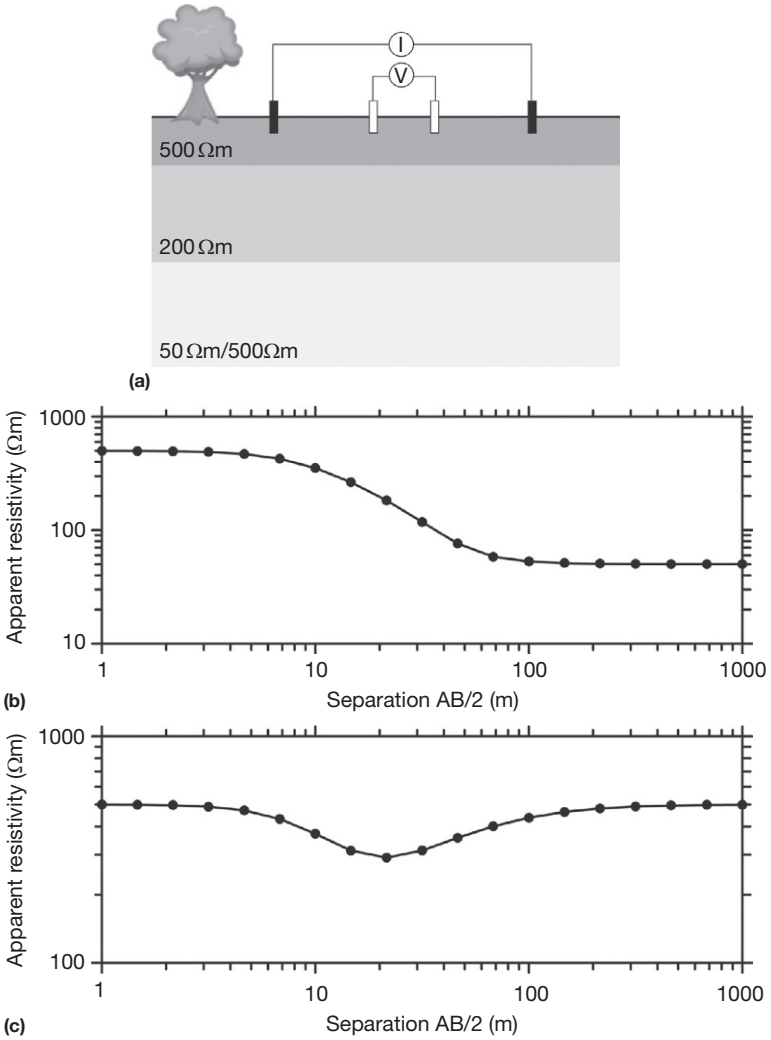


Figure 6 Three-layer Schlumberger resistivity sounding. (a) Resistivity variation. The thickness of the upper layer is 5 m; the thickness of the middle layer is 10 m. (b) Variation in apparent resistivity with electrode separation for three-layer model with the lower layer equal to 50 Ωm. (c) as in (b) but the lower layer is 500 Ωm.

technique, although measurements are more challenging and more labor-intensive (as discussed later). Four-electrode measurements are again made, which can be done in the time domain (as in DC resistivity) or in the frequency domain.

Although IP phenomena were reported much earlier (for earlier historical development, see Seigel at al. (2007)), it was not until the 1950s that the measurement approach began to develop, primarily driven by the needs of the mining and petroleum industries, with significant developments in North America and USSR. Bleil (1953) first coined the term 'induced polarization' in relation to mining geophysics. During much of the twentieth century, IP remained principally a mineral prospecting method (e.g., Komorov, 1972; Pelton et al., 1978); however, in the latter part of the century, interest in the use of this approach for 'environmental' applications developed (e.g., Bodmer et al., 1968; Olheoft, 1985; Olorunfemi and Griffiths, 1985; Roy and Elliot, 1980; Vacquier et al., 1957) with the realization that hydrogeologic and hydrochemical properties and states may be detectable using measurements of electrical

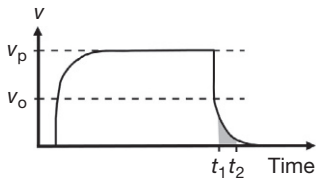


Figure 7 Overvoltage effect in time-domain measurement.

polarization. The last decade has seen a significant growth in related applications of IP (e.g., Flores Orozco et al., 2011; Kemna et al., 2004; Slater and Lesmes, 2002a).

11.08.2.2.1 Chargeability

Historically, IP was noted as an overvoltage effect. Figure 7 shows a typical measured voltage signal. When the current source is switched off, the voltage does not return to zero but drops suddenly from the primary voltage, V_p, to a voltage V_o

and then decays over time. The charge-up behavior is also reciprocated (as shown in Figure 7). In a DC resistivity measurement, this overvoltage effect can lead to spurious readings since we assume that the measured voltage is the primary voltage, V_p: if a significant IP effect exists, then V_p may not be reached if the time duration of current injection is too low (Figure 7). In time-domain IP measurements, this overvoltage effect is recorded, allowing some inference of the polarization properties of the subsurface.

In the time domain, the IP effect is measured as an apparent chargeability. Seigel (1959) defined the apparent chargeability, m_a, as the ratio of V_o to V_p (Schlumberger first labeled this term as the 'coefficient' in his 1939 patent). Such a definition has limited practical value as the secondary voltage V_o is difficult to measure. Consequently, other measurable quantities have been defined.

The integral chargeability, M_a, is the most widely used time-domain measure of the IP effect. This is defined as

$$M_a = \frac{1}{(t_2 - t_1)} \frac{1}{V_p} \int_{t_1}^{t_2} V(t)\,dt \qquad [12]$$

where t_1 to t_2 is the time interval of the measurement (see Figure 7). Although M_a is dimensionless, it is usually expressed in mV V^{-1}, that is, as a part per thousand (‰), although many early mineral prospecting applications reported chargeabilities as several %. M_a is sometimes expressed without the time interval denominator, in units of ms mV V^{-1} or ms. Given the potential ambiguity of such a measurement (the value depends on the sampling period), the Newmont standard (named after the Newmont Mining Corporation) was widely adopted in early IP applications. In the Newmont standard, the switch on and off times are 3 s and the integrating interval is 1 s (this is also referred to as M_{331}) (Sumner, 1976).

The integral in eqn [12] is achieved through sampling at discrete times (typically 10–20) over a user-defined time interval t_1 to t_2. Modern IP instruments allow these to be recorded and provide an integral measure (M_a). The selection of on/off time for the current transmitter is important in time-domain IP measurements: if the time interval is too short, then the signal will not have decayed fully before the next cycle (it is common to measure DC resistivity and IP over several cycles in order to improve accuracy and obtain some measure of repeatability). Consequently, superposition of overvoltage effects can result, which can lead to erroneous measurements (Fiandaca et al., 2012).

Given the fact that the overvoltage is much smaller than the primary voltage, the time-domain IP measurements can require much higher input current than in DC resistivity surveys and also more stacking of measurements (repeated cycles). IP measurements thus often require longer survey times than DC resistivity surveys, and the selection of array configuration can be important given the relatively low secondary voltages.

Since the DC apparent resistivity is also measured, time-domain measurements provide both resistivity and chargeability information. As argued by Slater and Lesmes (2002b), however, chargeability, M_a, can be an ambiguous measure since, as evident from eqn [12], it will be affected by the bulk resistivity, ρ. As a result, low resistivity, for example, due to high pore fluid conductivity, will suppress the measured chargeability even though the polarization along the fluid–solid interface may be significant. To overcome this, Slater and Lesmes (2002b) advocated the use of a normalized chargeability measure, defined as

$$M_{aN} = \frac{M_a}{\rho} = \sigma M_a \qquad [13]$$

Measurements in the time domain are not constrained to analysis using the chargeability concept, as discussed later in this section.

11.08.2.2.2 Complex resistivity/conductivity

In the frequency domain, IP phenomena can be represented as a complex electrical resistivity $\rho^*(\omega)$ (or conductivity, $\sigma^*(\omega)$) that varies with frequency (ω), which can be expressed as

$$\frac{1}{\rho^*(\omega)} = \sigma^*(\omega) = |\sigma(\omega)| e^{i\varphi(\omega)} = \sigma'(\omega) + i\sigma''(\omega) \qquad [14]$$

where * denotes a complex term, $|\sigma(\omega)|$ is the magnitude of conductivity, φ is the phase angle between injected current and measured voltage, $\sigma'(\omega)$ is the real component of conductivity, $\sigma''(\omega)$ is the imaginary component of conductivity, $\omega = 2\pi f$ is the angular frequency representation of frequency f, and $i = \sqrt{-1}$.

The measured phase angle, φ, is related to the real and imaginary components according to

$$\varphi = \tan^{-1}\left(\frac{\sigma''}{\sigma'}\right) \cong \left(\frac{\sigma''}{\sigma'}\right) \qquad [15]$$

where the approximation on the right-hand side of eqn [15] applies to small phase angles (<200 mrad). Note that the phase angle will be positive when expressed in terms of conductivity or negative when expressed in terms of resistivity.

In IP measurements, we are principally concerned with the polarization properties of the subsurface (see Glover, Chapter 11.04 for details of processes contributing to polarization in porous media). Like chargeability, the phase angle, φ, is ambiguous in this respect since, from eqn [15], it is a function of both in-phase and out-of-phase conductivity. And thus, like the necessary treatment of chargeability in eqn [14], it is preferable to focus interpretation of complex resistivity in terms of the imaginary conductivity, σ''.

Unlike chargeability, complex resistivity is a uniquely defined measurement quantity. It may be obtained using the same four-electrode arrangement as DC resistivity, with the same constraints noted for chargeability measurement. Frequency domain instruments will measure the ratio of the magnitude of the peak voltage to peak current and the angular phase difference between voltage and current, for a given frequency of current injected in a sinusoidal waveform.

The phase angle, φ, and chargeability are likely to exhibit some relationship. Kemna et al. (1997) illustrated how a theoretical link can be established based on the time-domain instrument settings. Alternatively, one may evaluate the relationship through empirical means, for example, Mwakanyamale et al. (2012). For most cases, a relationship of the form

$$\varphi \cong M \text{ to } 1.5M \qquad [16]$$

is likely, where the conductivity phase angle, φ, is measured in mrad and M is measured in mV V^{-1}. The usefulness of such a

relationship is that chargeability measurement from commonly available time-domain IP instruments can be utilized for modeling and interpretation in terms of complex resistivity (see Section 11.08.4).

Measurements of complex resistivity in the frequency domain can be made at a single frequency, f, for example, 1 Hz, or over a frequency range in order to assess spectral properties. Measurements of frequency-dependent electrical properties of geologic materials have revealed considerable information in such spectral measures over a 1 mHz to 1 kHz range (e.g., Lesmes and Friedman, 2005); this is often referred to as spectral IP. However, the measured frequency range is constrained in field deployment: low frequency measurements are limited by potentially long data acquisition times; high frequency measurements are limited by coupling effects. Consequently, field measurements will be constrained, at best, to only three frequency decades, for example, 0.1–100 Hz.

The shape of the spectrum of electrical resistivity measurements is sensitive to the distribution of timescales of the electrical polarization processes within the geologic material, which may, for example, be related to the grain-size distribution (e.g., Lesmes and Friedman, 2005). For a narrow range of polarization timescales, a distinct peak in the imaginary conductivity spectrum will be seen, provided sufficient frequency range is used. In contrast, relatively flat spectra will be observed for media exhibiting a broad range of polarization timescales.

Spectral models may be utilized to aid analysis and interpretation of spectral data. For relatively flat phase angle spectra, the constant phase angle (CPA) model is often used. Examples of observed CPA behavior from field investigations include Van Voorhis et al. (1973) and Börner et al. (1993). In this model, the complex conductivity is expressed as (Börner and Schön, 1991)

$$\rho^*(\omega) = |\rho|_n \left(\frac{\omega_0}{\mathrm{i}\omega}\right)^p \text{ or } \sigma^*(\omega) = |\sigma|_n \left(\frac{\mathrm{i}\omega}{\omega_0}\right)^p \qquad [17]$$

where $|\rho|_n$ is the magnitude of the complex resistivity and some arbitrary frequency, ω_0. In the conductivity equivalent, $|\sigma|_n$ is the magnitude of the complex conductivity at ω_0. In eqn [17], p is a power law dispersion model parameter given by $2\varphi/\pi$ (φ expressed in terms of conductivity phase angle). The exponent p describes the frequency dependence of both real and imaginary components of complex resistivity; its value has been linked to the geometric structure of the chargeable surface (see discussion by Kemna (2000)).

Where a distinct peak in the phase angle spectra is observed, the Cole–Cole model (after Cole and Cole, 1941) has been widely used. The Cole–Cole model represents the cumulative effect of a distribution of the Debye relaxations. Although Cole and Cole's original model was developed in order to explain dielectric phenomena, its functional form has been used to represent resistivity spectra following the work of Pelton et al. (1978), who expressed the spectral model as

$$\rho^*(\omega) = \rho_\infty + \frac{\rho_0 - \rho_\infty}{1 + (\mathrm{i}\omega\tau)^c} \qquad [18]$$

where ρ_0 and ρ_∞ are the low- and high-frequency asymptotes of the resistivity magnitude, τ is a time constant (sometimes referred to as a relaxation time, characterizing the position of the peak in phase angle spectra), and c is an exponent parameter, which is inversely related to the width of the dispersion shown in the spectra.

If we follow the definition of intrinsic chargeability (note that this is different to the measured apparent chargeability used earlier)

$$m = \frac{\rho_0 - \rho_\infty}{\rho_0} \qquad [19]$$

then we can write the Pelton Cole–Cole model as

$$\rho^*(\omega) = \rho_0 \left[1 - m\left(\frac{\rho_0 - \rho_\infty}{1 + (\mathrm{i}\omega\tau)^c}\right)\right] \qquad [20]$$

Equation [20], expressed in terms of conductivity, is (e.g., Jones, 2002)

$$\sigma^*(\omega) = \frac{1}{\rho_0}\left[1 + m\left(\frac{(\mathrm{i}\omega\tau)^c}{1 + (\mathrm{i}\omega\tau)^c(1-m)}\right)\right] \qquad [21]$$

In the Cole–Cole model, a single distinct peak phase angle is observed at the frequency (Major and Silic, 1981):

$$\omega_{\mathrm{pk}} = \frac{1}{\tau(1-m)^{0.5c}} \qquad [22]$$

Complex resistivity measurements may not indicate monomodal spectra, as in the Cole–Cole model: high frequency measurements are prone to interference from inductive and capacitive coupling. Although not a physically based representation of the effect, success has been shown in using a superposition of multiple Cole–Cole relaxation models in order to 'filter' the coupling effect (e.g., Kemna, 2000; Pelton et al., 1984). The approach of Kemna (2000) leads to the reformulation of eqn [20] as

$$\rho^*(\omega) = \rho_0 \left(1 - \sum_{i=1}^{N}\left[m_i\left(1 - \frac{1}{1 + (\mathrm{i}\omega\tau_i)^{c_i}}\right)\right]\right) \qquad [23]$$

where $i = 1, 2, \ldots, N$ is the relaxation model with parameters m_i, c_i, and τ_i.

Although the most widely used, the Cole–Cole model is not the only spectral model used for frequency analysis of IP data. Dias (2000), for example, listed several alternative similar forms; other formulations include those based on decomposition of individual Debye relaxations (e.g., Keery et al., 2012; Morgan and Lesmes, 1994; Nordsiek and Weller, 2008).

While true frequency-domain measurements of complex resistivity are relatively common in laboratory-scale instruments (and widely used in the general field of electrical impedance spectroscopy, e.g., Barsoukov and Macdonald, 2005), they are less common in field-based instrumentation. However, the transformation to the frequency domain can be achieved by Fourier analysis of the time-domain signals due to a pulsed square waveform, as illustrated, for example, by Kemna (2000), although the frequency range will be limited by the sampling interval of the time-domain measurements.

Rather than measuring an integration of the polarization effect in the time domain, attempts have been made to recover spectral information, often in terms of the four Cole–Cole parameters, using measurements of the voltage decay (e.g., Johnson, 1984; Yuval and Oldenburg, 1997). As noted by Fiandaca et al. (2012), modeling of the superposition of input series is necessary if the time duration of current switch

off is relatively short, since the decaying voltage from one cycle may be present during the next cycle. It is also important to note that instruments will apply low-pass (e.g., 10 Hz) filters to measured responses in order to remove noise (e.g., 50 Hz mains power effects) and attempts to utilize the discrete measured voltage decay need to account for such filters.

11.08.2.3 Measurement Errors

DC resistivity and IP measurements in the field can suffer from a range of sources of error, which, if not addressed correctly (as discussed in **Section 11.08.4.2**), can have a significant impact on the interpretation of survey results. High contact resistance between the measurement electrodes and the ground can be particularly problematic, although most modern instruments provide estimates of contact resistance prior to a survey and reduction of the contact resistance is relatively easy to achieve in ground surface surveys by the addition of small quantities of saline fluid around the electrodes. It is also important to take into account the receiver voltage levels in a given measurement: high geometric factors, combined with low input voltage, can lead to voltages that are close to instrument resolution. Natural self-potentials also need to be accounted for, particularly if they are not stable over time, although most modern instruments incorporate appropriate filters to reduce such effects.

IP measurements are particularly prone to measurement errors, and in comparison to DC resistivity, much more care is needed in order to reduce such effects. Since the IP measurement, in the time domain, is of a secondary voltage, then the magnitude of the signal can be even more problematic than in DC resistivity surveys (see, e.g., Gazoty et al, 2013). As a result, much greater input voltages are necessary for IP measurements in the field, particularly for large dipole separations. In addition, given the original mineral exploration application of IP, traditionally, IP surveys have been conducted over large length scales, requiring large electrode spacing and large current injection. On the grounds of health and safety, separating the active area associated with transmitter electrodes from the measurement region has clear advantages if excessive current injection is required.

Traditionally, IP nonpolarizing receiver electrodes (e.g., copper–copper sulfate and lead–lead chloride) have been used, adding greater labor costs to surveys. However, as demonstrated by Dahlin et al. (2002), conventional stainless steel electrodes, as used in DC resistivity surveys, can be utilized with appropriate correction (in some cases built into the measurement system by instrument manufacturers). LaBrecque and Daily (2008) provided some useful comparisons of electrode materials for IP surveys.

IP measurements can be made in the field with similar four-electrode configurations to those used in DC resistivity. One constraint, however, is the capacitive and inductive coupling that may occur (see Sumner, 1976). Consequently, dipole–dipole electrode configurations have been favored since they allow some electrical 'isolation' of transmitter and receiver cable pairs. Dahlin et al. (2002) had demonstrated, however, that multicore cables, widely used for DC resistivity surveys, can prove effective for environmental-type applications of IP.

Negative apparent resistivity values in DC resistivity surveys often highlight problems with contact resistance and signal strength. In contrast, apparent negative IP effects can result from contrasts in subsurface electrical properties (e.g., Martinho and Almeida, 2006; Nabighian and Elliot, 1976), making interpretation of raw IP measurements challenging. Such measurements may be assessed to be erroneous; however, through appropriate modeling (see **Section 11.08.4.2**), these measurements provide information about the subsurface and should not be rejected on the grounds of polarity alone.

The quality of DC resistivity and IP measurements is often determined through repeatability, that is, by assessing the variability of responses from multiple injected cycles. While these are useful direct in-field indicators of data quality, sources of error may not be random and could, in theory, repeat. An alternative measure of data quality is reciprocity (e.g., Parasnis, 1988). Measurements made after switching injection and receiver pairs should be identical; the difference is often termed a reciprocal error and can often be a much better assessment of data quality in DC resistivity and IP surveys (see, e.g., Slater and Binley, 2006; Slater et al., 2000). Reciprocity checks, however, require double the normal set of measurements, and for this reason, several surveys are conducted without such checks (e.g., Dahlin and Leroux, 2012), particularly when one considers that some modern multichannel instruments do not permit full flexibility of multichannel configurations, and thus, conducting reciprocal measurements with automated systems may take longer than a normal survey. Furthermore, one has to recognize that charging of transmitter electrodes in a four-electrode measurement may render the transmitter electrodes unsuitable as receiver electrodes until fully depolarized (e.g., Dahlin, 2000), which may be particularly problematic in IP surveys where such low voltages are recorded.

11.08.3 Field Configuration

DC resistivity and IP surveys can be conducted in a wide range of configurations, the choice of which will depend upon (1) objectives of the survey; (2) site access constraints, including the availability of electrode sites, for example, boreholes; (3) instrument and labor constraints; and (4) health and safety issues, for example, potential impact of surveys on site and associated personnel, livestock, etc. Most surveys are conducted using electrode arrays placed on the ground surface for either 1-D or 2-D (vertical) analysis (although the growing availability of multichannel computer-controlled instrumentation and efficient data processing tools has led to a growth in the use of 3-D investigations). Traditionally, DC resistivity and IP surveys were conducted using surface-based electrode arrays, and data processing tools were constrained to 'type curve' analysis or pseudosections (see **Sections 11.08.3.1** and **11.08.3.2**); however, the availability of generalized modeling tools now permits full flexibility in electrode configuration, not just using surface-based electrodes. In this section, we discuss some common electrode configurations for field surveys before presenting the data modeling approaches in the following section. We illustrate the approaches using DC resistivity; however, these are all directly applicable to IP

applications, using the chargeability or complex resistivity formulation.

11.08.3.1 1-D Electrical Sounding

In a vertical electrical sounding (VES), apparent resistivity measurements are made at different electrode spacings, centered about a common point. As the electrode array size increases, we 'sound' to greater depths, as shown in the earlier examples. The Schlumberger array is commonly used for VES, keeping the potential electrode dipole (M and N) fixed and moving the current electrode dipole (A and B). This is a popular choice because the fixed potential electrode dipole means that not only the measurements are relatively insensitive to lateral variation in resistivity but also, from a practical standpoint, surveys can be completed more efficiently as only one of the dipoles is moved for each measurement. The disadvantage is that for small MN spacings, the signals (potential differences) can be relatively weak and thus high-sensitivity instrumentation may be required.

VES surveys require just four electrodes, each with a suitable cable to connect to the instrument. For a Schlumberger array VES, the spacing AB is normally increased in a logarithmic sequence (i.e., the logarithm of the spacing AB is increased linearly in each step). The results are often presented as a plot of the logarithm of apparent resistivity versus the logarithm of AB/2, as shown in **Figure 5(c)**. Using data modeling tools (see **Section 11.08.4.1**), these data are analyzed in order to recover a 1-D resistivity structure (defined as a series of layer thicknesses and associated resistivity) that is consistent with the measured response.

11.08.3.2 2-D Surface-Based Imaging

If we extend the 1-D electrical sounding concept to a series of soundings centered along a transect, then we are able to develop a 2-D vertical profile. Such surveys can be conducted with only four electrodes (and associated cables), although this can be extremely labor-intensive; modern multielectrode instruments now permit connection to an array of electrodes, connected to the instrument through a multicore (or series of multicore) cable(s). Griffiths and Turnbull (1985) outlined one of the earliest field geophysics prototypes, using an 8-core cable, although similar developments in electrical imaging were advancing at the same time in other fields (Lytle and Dines, 1978, 1980; Wexler et al., 1985).

The choice of quadrapole configuration will depend on a number of factors: (1) the objective of the survey (e.g., if the survey is to delineate lateral variability, then a dipole–dipole configuration may be suitable); (2) signal-to-noise constraints, which will depend on subsurface geoelectrical properties and the instrumentation available; (3) survey time constraints; (4) whether IP measurements are required; and (5) physical site constraints (if applicable).

Data from 2-D imaging surveys are normally presented as a pseudosection of apparent resistivity (or chargeability or complex resistivity). **Figure 8** illustrates how a pseudosection is constructed, in this case for a dipole–dipole configuration, with each transmitter or receiver dipole width equal to 2 m. As we increase the separation of the dipoles (n in eqn [10]), the depth of sensitivity increases; this is graphically shown in the pseudosection as a level. The apparent resistivity (e.g., from eqn [10] for the dipole–dipole configuration), or the apparent chargeability or complex resistivity, is then computed and assigned to this position. In some cases, the level axis is shown as a pseudodepth using an appropriate assignment of a depth of maximum sensitivity. However, it should be apparent from **Figure 2** that a single depth may be misleading. It is also important to note that the pseudosection does not reveal the resistivity structure; it is simply a means of displaying data. We illustrate this in the succeeding text with a model example. Prior to the availability of data inversion tools (see later), interpretation of field data relied on analysis of pseudosections (e.g., Edwards, 1977); however, modeling of field data is now easily achieved, dismissing the need for such treatment of pseudosections.

In **Figure 9**, a 2-D structure of resistivity is shown. A 1 m thick, 100 Ωm layer overlies a 200 Ωm formation, in which a 5 m wide, 500 Ωm unit protrudes vertically. Twenty-five electrodes are placed on the ground surface at 2 m spacing. Note that although this example is used for illustration only, for investigation scales (e.g., electrode spacing) larger/smaller, the diagram can be reduced/increased in scale. **Figure 10** shows pseudosections for dipole–dipole and Wenner configurations, based on the resistivity structure in **Figure 9**. These were computed using a solution to eqn [1] for the given resistivity distribution, which is assumed to extend in the strike direction. There are several points worthy of note. First, the dipole–dipole layout offers potentially greater depth of investigation at the left and right margins of the survey line (however, we will return to depth of investigation in **Section 11.08.4.2**). Second, the pattern of apparent resistivity

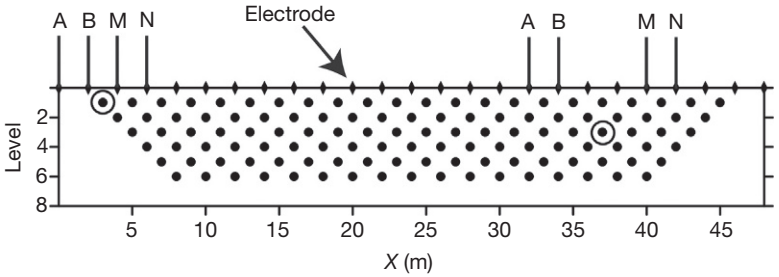


Figure 8 2-D surface imaging pseudosection construction schematic showing locations of assigned measurements for two dipole–dipole configurations. The solid symbols indicate the entries for the entire pseudosection. The two circled symbols show assigned locations for the two measurements shown.

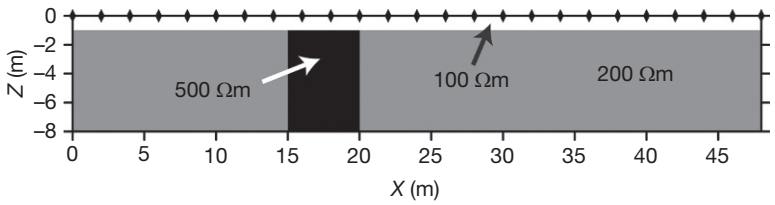


Figure 9 Resistivity structure, for example, model. Twenty-five electrodes at 2 m spacing are located on the ground surface (indicated by diamond symbols).

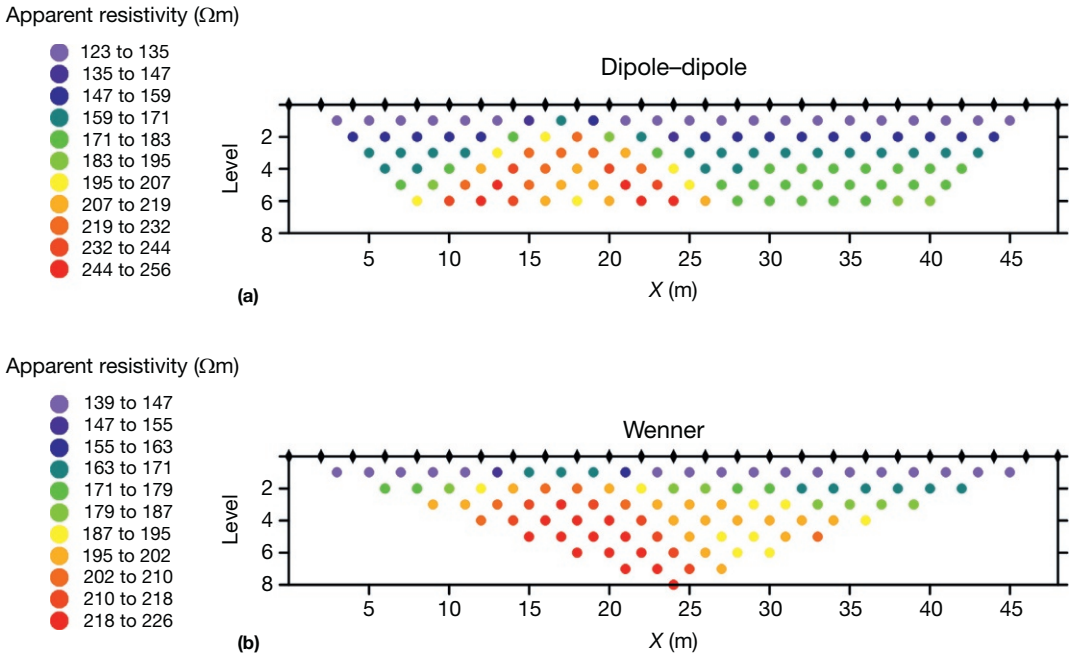


Figure 10 Pseudosections for resistivity model in **Figure 9**. (a) Dipole–dipole configuration; (b) Wenner configuration.

in the pseudosection differs for the two geometries, again highlighting that these images do not represent an image of the subsurface. Third, the values of apparent resistivity differ for the two geometries, as a result of different sensitivity patterns (e.g., **Figure 2**). In **Section 11.08.4.2**, we will examine methods to process these 2-D imaging data and compare interpreted resistivity models for the two cases shown in **Figure 10**.

In the aforementioned example, if we were to take the depth of investigation estimates based on the analysis of Roy and Apparao (1971), the first three levels in the Wenner configuration would translate to depths of 0.7, 1.3, and 2.0 m, and in the dipole–dipole configuration, the depths would be 1.1, 1.4, and 1.8 m.

As stated earlier, 2-D imaging surveys are now routinely performed with multielectrode configurable instruments, coupled with multicore electrode cables. Cables with 25 electrode wires, with electrode takeouts at 5 m spacing, are manageable by one field operator (coupled with a cable drum, a single cable will be typically 15–20 kg). Placing the instrument in between two such cables permits a survey line of 245 m with 5 m electrode spacing. Longer cable lengths (more electrodes or longer electrode spreads) may require transporting/lifting gear to aid deployment. However, relatively long survey lines can be achieved using 'roll along' techniques (e.g., Dahlin, 1996); in this mode, the instrument is moved to the end of the survey line and the cable furthest away moved to form the extension of the transect.

Although most 2-D DC resistivity surveys are conducted using manual placement of electrodes/cables, as described &INS id="ins11.08. orig="above";in the preceding text, there has been some interest in the use of towed electrode arrays for more high (lateral) resolution and rapid coverage of longer transects. Such surveys are often referred to as continuous vertical electrical soundings (CVESs) since most analysis treats datasets as a series of 1-D soundings. Panissod et al. (1998) and Christensen and Sørensen (2001) illustrated different approaches using vehicle towed arrays, which have proved effective, particularly for relatively shallow investigations in archeology (e.g., Simpson et al., 2010) and agronomy (e.g., André et al., 2012).

CVES surveys can also be conducted using waterborne arrays in order to map variation in bed geoelectrical properties. Several manufacturers now offer floating electrode cable accessories. Examples of such surveys using DC resistivity include Bradbury and Taylor (1984), Allen and Merrick (2007), Sambuelli et al. (2011), and Rucker et al. (2011). In the last example, Rucker et al. (2011) completed over 660 line kilometers along the Panama Canal, using a 170 m long, 11-electrode floating

array, delivering a maximum of eight dipole–dipole measurements at each location, at 3.75 m intervals along the transects. Most continuous electrical surveys have focused on DC resistivity, given the need for a reasonable acquisition time for reliable IP measurements. However, Slater et al. (2010) demonstrated the effectiveness of waterborne continuous DC resistivity and IP soundings in their survey of a reach of the Columbia River, USA.

11.08.3.3 3-D Surface-Based Imaging

2-D surface-based imaging has naturally progressed to 3-D investigations as multielectrode instruments, with multichannel capability, have advanced, coupled with the recent computational (hardware and software) developments. Quasi 3-D imaging is easily achievable through the use of multiple 2-D imaging transects (e.g., Cassiani et al., 2006; Dahlin and Loke, 1997); however, such a configuration has limited sensitivity normal to each transect. True 3-D configurations are now utilized for surface arrays, although data collection time can be a significant constraint, and given the required number of quadrapole configurations, such surveys are often limited to relatively small-scale investigations. Loke and Barker (1996a) and Nyquist and Roth (2005), among others, discuss possible surface electrode array configurations for 3-D imaging. Park and Van (1991) describe one of the earliest applications of 3-D resistivity imaging using surface-based arrays, in this case using a pole–pole configuration. We revisit 3-D imaging later in the chapter, when addressing modeling and inversion.

11.08.3.4 Borehole Logging

DC resistivity measurements are routinely used in geophysical logging of boreholes and can greatly assist in, for example, the delineation of hydrogeologic units (e.g., Keys, 1990), particularly when coupled with other well log data, for example, natural gamma. Electrodes are installed on a borehole sonde, connected to the logging unit on the surface. As the sonde is slowly lowered in a water-filled, uncased, borehole, apparent resistivity measurements are made. The normal configurations used are short normal, long normal, and lateral, each of which has a different four-electrode configuration to permit either greater resolution along the borehole or greater sensitivity into the format (as the expense of reduced vertical resolution). It is normal to conduct well logs with at least two configurations to aid interpretation. For more details, the reader is referred to Telford et al. (1990).

11.08.3.5 Borehole-Based Imaging

Although most DC resistivity and IP imaging surveys are conducted using electrode arrays located on the ground surface, boreholes are frequently used for electrode sites, permitting greater focus at depth. For borehole electrode placement, all that is required is electrical contact with the formation. This can be achieved in open holes (or fully screened plastic-cased holes) beneath the water table; in such a configuration, the electrode arrays only need to be placed for the duration of the survey, and several instrument manufacturers now offer borehole electrode cables as accessories. In partially plastic-cased boreholes, electrodes need to be installed within the borehole annulus using a suitable backfill to ensure adequate electrical contact with the formation. The same procedure applies to borehole installations above the water table.

When electrodes are placed within the borehole water column, consideration must be made for the effect of short-circuiting of current up/down the water column. Studies have demonstrated the impact of neglecting such effects on resultant images (Doetsch et al., 2010; Nimmer et al., 2008; Osiensky et al., 2004). Some field investigations have attempted the use of electrical isolators along the water column. However, with suitable modeling tools, the effects of the water column can be accounted for, as discussed later.

Electrodes in boreholes may be utilized in a number of ways, as shown in **Figure 11**. In the mise-à-la-masse method, the current source is placed in a borehole at depth, while the other current electrode is located on the surface some distance away from the borehole. Potential measurements are then made on the ground surface relative to a remote electrode. The pattern of pole–pole apparent resistivities is then easily displayed, which may help explain its early popularity. However, the interpretation of results can be challenging without suitable modeling tools and sensitivity to geoelectrical contrasts can be limited. Examples of mise-à-la-masse application areas include mineral exploration (e.g., Bhattacharya et al., 2001; Mansinha and Mwenifumbo, 1983), geothermal application (e.g., Beasley and Ward, 1986), and hydrogeology (Bevc and Morrison, 1991; Nimmer and Osiensky, 2001).

Electrical measurements along a single borehole can be used to develop a profile of resistivity, as in a borehole well log, albeit with significantly reduced vertical resolution due to the small number of measurement locations along the borehole (e.g., Binley et al., 2002a). By expanding the electrode dipole lengths, one is able to sense further into the formation (as in a surface-deployed survey), although in this case, the volume of investigation is radially symmetrical about the borehole (e.g., Tsourlos et al., 2003).

As in surface measurements, pseudosections can be computed, although apparent resistivity calculations must account for the noninfinite surface boundary (expect for very deep placement). For an infinitely deep electrode, eqn [3] becomes

$$V = \frac{I\rho}{4\pi r} \qquad [24]$$

However, if the current source electrode, A, is placed at depth z_A ($z_A > 0$) and voltage measured with electrode at depth z_M, directly above (or below) A, then accounting for the surface boundary (at $z = 0$) results in

$$V = \frac{I\rho}{4\pi}\left[\frac{1}{|z_A - z_M|} + \frac{1}{z_A + z_M}\right] \qquad [25]$$

Consequently, the geometric factor for the general AB–MN four-electrode measurement is

$$k = 4\pi\left[\left(\frac{1}{|z_A - z_M|} + \frac{1}{z_A + z_M}\right) - \left(\frac{1}{|z_B - z_M|} + \frac{1}{z_B + z_M}\right) - \left(\frac{1}{|z_A - z_N|} + \frac{1}{z_A + z_N}\right) + \left(\frac{1}{|z_{B-} z_N|} + \frac{1}{z_B + z_N}\right)\right]^{-1} \qquad [26]$$

where z_B and z_N are the depths of electrodes B and N, respectively.

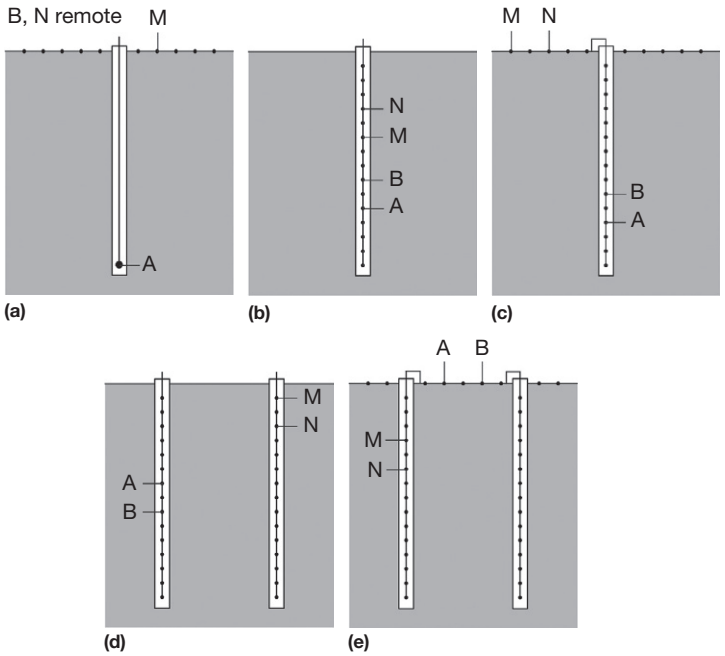


Figure 11 Schematic of possible uses of boreholes for electrical imaging. (a) The mise-à-la-masse, (b) single borehole, (c) borehole-to-surface, (d) cross borehole, and (e) cross borehole with surface.

Single-borehole electrode arrays may be supplemented with surface electrodes (Marescot et al., 2002; Tsourlos et al., 2011); however, poor signal-to-noise can result from deep electrode placement with short dipole spacing. Borehole-to-surface surveys could, however, be advantageous when used with varying surface azimuthal arrays in order to assess anisotropy, for example, in fractured environments. In addition, the interpretation of borehole-to-surface array data can be challenging due to the more complex sensitivity patterns (in comparison with a single surface or borehole linear array). Although apparent resistivities are easily computed (with eqn [26]), graphic presentation of a pseudosection is difficult because of the sensitivity distribution.

Cross borehole surveys employ two or more boreholes as electrode sites, which can be supplemented with surface electrodes. Such a configuration permits high-resolution 2-D or 3-D imaging at a targeted depth and also removes the need for access to appropriate surface sites for electrodes (which may be a constraint in industrial sites, e.g., or for investigations under buildings, underground storage tanks, etc.). Measurements can be made in different electrode configurations (**Figure 12**). Bing and Greenhalgh (2000), among others, have explored different dipole patterns on sensitivity. Universally optimum schemes are unlikely to exist since performance will depend on (1) spacing of boreholes, (2) subsurface geoelectrical properties, (3) availability of supplementary surface electrode sites, and (4) depth of borehole electrode array. We will address this in the next section.

Clearly, the sensitivity pattern of the measurements will depend on the spacing of the boreholes; irrespective of the dipole configuration used, high sensitivity will exist close to the boreholes and will be low mid-way between the boreholes. Consequently, cross borehole imaging is constrained to imaging relatively small survey areas; typically, the distance between boreholes is less than the shortest length of an electrode array in a borehole.

Daniels (1977) outlined the concept of borehole-based DC resistivity and IP measurements; however, since a pseudosection is not applicable in this configuration, applications did not develop until the 1990s because of the need for more sophisticated modeling tools (early works in this area being: LaBrecque and Ward, 1990; Sasaki, 1989; Shima, 1990, 1992). Daily and Owen (1991) is one of the earliest applications of cross borehole resistivity imaging, which arose out of pioneering concepts by Lytle and Dines (1980). The Lawrence Livermore group continued with further demonstration of cross borehole resistivity imaging in a wide range of applications (e.g., Daily et al., 1992; Ramirez et al., 1993); see also Daily et al. (2004a). Many other development and applications followed in the 1990s (e.g., Morelli and LaBrecque, 1996; Schima et al., 1996; Slater et al., 1996; Slater et al., 1997a,b; Zhou and Greenhalgh, 1997), all of which were based on 2-D configurations. 3-D cross borehole imaging applications are now relatively common (e.g., Binley et al., 2002b; Doetsch et al., 2012a; Wilkinson et al., 2006a). Following successful application of cross borehole imaging to DC resistivity problems, extensions to IP followed (e.g., Kemna et al., 2004). The term electrical resistivity tomography (or electrical resistance tomography) was adopted as a term for cross borehole electrical imaging, although this term is now widely used for describing 2-D and 3-D surface imaging. Other terms have also been used to describe electrical resistivity imaging, for example, electrical impedance tomography was adopted from the process tomography and biomedical imaging fields to refer to complex resistivity imaging, although in these other fields, it is used broadly to cover the real component of resistivity (see, e.g., Webster, 1990); microprocessor-controlled resistivity traversing (Griffiths et al., 1990).

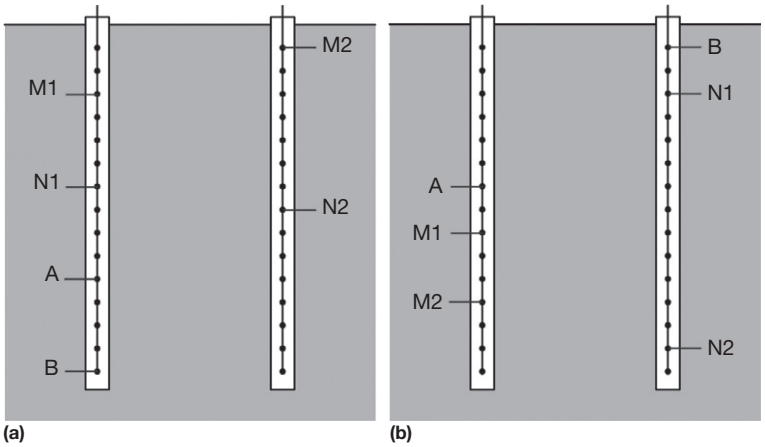


Figure 12 Cross borehole geoelectrical measurement configurations. (a) AB–MN, (b) AM–BN. Two example measurements in each scheme are shown with electrodes A, B, M1, N1 and A, B, M2, N2.

The use of metal-cased boreholes for electrical imaging was first investigated by Schenkel (1995) (see also Schenkel and Morrison, 1990). Ramirez et al. (1996) took this concept further by attempting to use metal-cased boreholes as individual long electrodes, thus providing a horizontal image of electrical resistivity, integrated over the borehole length. Rucker et al. (2010) have more recently adopted the same concept.

Borehole-based imaging concepts have also been adopted for mine/tunnel investigations. Sasaki and Matsuo (1993) described an application of surface to tunnel imaging of a copper skarn mine; Kruschwitz and Yaramanci (2004) illustrated how electrical imaging can be used to map fracture patterns in a rock salt mine; Van Schoor and Binley (2010) showed how tunnel-to-tunnel imaging can be used to detect potholes in a South African platinum mine.

11.08.3.6 Optimal Measurement Configuration

As outlined earlier, for surface-based electrode placement, different measurement configurations (e.g., Wenner and dipole–dipole) have advantages and disadvantages (see also Dahlin and Zhou, 2004). The choice of measurements will depend on several factors, including the geoelectrical structure of the target, data collection time constraints, and measurement signal-to-noise. It should also be noted that the traditional approach of adopting one specific configuration for a survey is no longer appropriate given that modern computer-controlled instruments coupled to multielectrode units reduce labor costs significantly and data inversion tools (covered later) should not be configuration-specific.

It will be apparent that with a wide range of geometric configurations of electrodes (surface, borehole, borehole–borehole, and borehole-to-surface), establishing a universally optimal set of measurements is impossible. For N electrode sites, there are $N(N-1)$ possible pole–pole measurements, although only half of these are independent because of reciprocity in the configuration. Any quadrapole measurement can, in theory, be computed from these $N(N-1)/2$ measurements by superposition, although it is not necessarily practical to do this in the field because of compounding errors in the superposition. Furthermore, many modern resistivity instruments with multichannel capability do not offer full flexibility in multichannel configuration (see, e.g., Stummer et al., 2002), and thus, data collection time for a complete set of independent measurements is a significant constraint.

In their pioneering concept on the impedance camera, Lytle and Dines (1978) recognized the need to determine optimal electrical measurement configurations. Attempts have been made to develop approaches for such optimum sets, for example, based on sensitivities to changes in resistivity (e.g., Furman et al., 2004). Perhaps, the most promising approach is that based on a sequential experimental design, in which an initial measurement set is used to derive an image of the subsurface geoelectrical properties, and then measurements from further configurations are added in order to enhance to resolution (see Stummer et al., 2004; Wilkinson et al., 2006b). Blome et al. (2011) and Wilkinson et al. (2012) extended these approaches, most notably by accounting for measurement errors in their analysis, along with the recognition that time delays are necessary between measurements of potential on an electrode that has previously been used for current injection (because of polarization). These approaches rely on the computation of the model resolution matrix (defined later), which can be computationally demanding. Despite relatively easy adaptation to parallel computation (e.g., Loke et al., 2010), these approaches are currently limited to 2-D imaging, and attempts to address IP imaging have, thus far, been ignored.

11.08.3.7 Monitoring and Automated Data Collection

Electrical methods are well suited to studying time-dependent subsurface processes: changes in electrical properties over time can give some insight into key hydrologic and hydrochemical processes. Many in situ groundwater remediation techniques, for example, result in changes in bulk electrical conductivity, and consequently, monitoring such processes with electrical methods may give insight into the effectiveness of the applied cleanup technology (e.g., LaBrecque et al., 1996a; Versteeg and Johnson, 2008). However, labor costs can restrict such monitoring, particularly for remote installations and long-term

studies. Recognizing this, a number of studies have developed automated data collection systems. Daily et al. (2004b) and LaBrecque et al. (2004) reported on applications of customized measurement systems, designed for remote data acquisition and data processing. In these two examples, electrode arrays were installed in boreholes, within a secured (restricted access) site, and thus, safety issues related to active electrodes were relatively straightforward to manage. This type of monitoring is often referred to as 'autonomous,' although it is really automated, since the system is typically not designed to use specific environmental (or other) criteria to control the measurement. More recent examples of such automated systems are reported in Ogilvy et al. (2009), for long-term monitoring of seawater intrusion in a coastal aquifer, and Wilkinson et al. (2011), for a number of examples, most notably landslide monitoring. Many instrument manufacturers now offer the capability of remote data acquisition as a standard feature.

11.08.4 Modeling and Inversion

As demonstrated earlier in this chapter, measured apparent resistivity (and chargeability) values do not provide directly a model of the subsurface geoelectrical properties. Interpretation of measured data thus relies on appropriate modeling tools. Forward models compute the apparent resistivity (or chargeability) response for a given geoelectrical structure of the subsurface. Inverse models determine the geoelectrical structure for a given set of measurements. Traditionally, 1-D sounding data were analyzed by manual curve matching procedures, using master and auxiliary curves for precomputed responses for different two- and three-layer models (see, e.g., Reynolds, 2011) and, in some cases, 3-D bodies (Dieter et al., 1969); that is, the inverse modeling step was not automated. Advances in robust inverse models for geoelectrical problems are one of the key factors that have contributed to the widespread use of these methods. We present some of the main developments here, but first, consider the forward modeling step, since this is a critical component of any inversion process (manual or automatic).

11.08.4.1 Forward Modeling

The solution of eqn [1] for a horizontally layered earth model, expressed in terms of apparent resistivity in a Schlumberger configuration, is given by the Stefanescu integral (e.g., Koefoed, 1979) of the form

$$\rho_A\left(\frac{AB}{2}\right) = \left(\frac{AB}{2}\right)^2 \int_0^\infty T(\lambda) J_1\left(\lambda \frac{AB}{2}\right) \lambda \, d\lambda \qquad [27]$$

where J_1 is the Bessel function of order one and the kernel $T(\lambda)$ is a function of the given layer thicknesses and resistivities. For a 1-D problem (which is considered in a VES), the integral in the equation in the preceding text can be readily calculated, for instance, using fast filtering techniques (e.g., Anderson, 1979).

If the resistivity distribution is considered 2-D (x–z), as shown by Hohmann (1988), eqn [1] can be written as

$$\frac{\partial}{\partial x}\left(\frac{1}{\rho}\frac{\partial v}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{1}{\rho}\frac{\partial v}{\partial z}\right) - \frac{vk^2}{\rho} = -\frac{I}{2}\delta(x)\delta(z) \qquad [28]$$

where

$$v(x, k, z) = \int_0^\infty V(x, y, z)\cos(ky)\,dy \qquad [29]$$

and k is the wave number. Once eqn [28] is solved for the transformed potential v, an inverse Fourier transform is applied to obtain V (e.g., Kemna, 2000).

For generalized 2-D and 3-D resistivity structures (i.e., problems requiring the solution of eqns [28] or [1], respectively), semianalytic solutions, such as eqn [27], are not available and numerical approximations must be determined. These are achieved through grid-based methods, commonly either finite difference or finite element-based. With these methods, the potential field (i.e., voltages, for a given current injection) is computed at node points (or cell centers), which are connected together to form a grid or mesh of cells or elements. Finite difference methods in applied mathematics date back to the early twentieth century (or earlier) and are based on differential approximations to the partial derivatives. Finite element methods are more recent (they emerged in the 1960s in the fields of civil and aeronautical engineering) and are based on integral approximations using variational calculus. Finite element methods do not necessarily provide greater accuracy (although variants of the approach can be exploited to enhance accuracy) than finite difference methods; the key advantages are the ability to geometrically represent arbitrary structures and mesh refinement in unstructured meshes. Finite difference methods have the advantage of being simpler to implement and can be computationally more efficient for specific problems, particularly in terms of computer storage.

In the 1970s, finite difference solutions to geoelectrical problems emerged, and in 1979, Dey and Morrison presented a 3-D finite difference model for the solution of the DC resistivity forward model (Dey and Morrison, 1979), with example applications to surface and borehole-to-surface problems. Applications of their model at the time relied on access to large supercomputer facilities, requiring significant resources to compute solutions to problems with 10 000 unknowns, which were then considered large. Dey and Morrison's code was widely used in many 2-D and 3-D applications that followed. Park and Fitterman (1990) described one such application in which data were collected on a grid of surface dipoles, thus requiring 3-D analysis. Park and Fitterman (1990) modeled their 3-D volume using $27 \times 21 \times 10$ finite difference cells, which required 4 Mb of computer core storage (equivalent to the term RAM on a modern computer) and was their storage limit. Furthermore, manual matching of the data to computed responses required several months' effort (Park and Van, 1991).

Modeling potential fields on structured grids (which are required for finite difference solutions) can lead to relatively high modeling errors because of constraints in discretization close to current sources. Lowry et al. (1989) outlined a scheme for removal of singularities in the potential field, which can significantly improve model estimates without requiring computationally constraining fine numerical grids. Kemna (2000) presented an extension of this for the general complex resistivity formulation.

Coggon (1971) formulated a finite element solution of the 2-D resistivity problem; later, Pridmore et al. (1981) extended this to 3-D. Although the solution of Pridmore et al. (1981) was based on tetrahedral finite elements (thus permitting the

development of unstructured meshes), they applied their algorithm to hexahedral (brick) elements, thus benefitting little in comparison with the finite difference solution of Dey and Morrison (1979). Pridmore et al. (1981) also highlighted the computational constraints of their modeling approach and finished with the remark about complex 3-D modeling: "In all probability, an efficient solution to this class of problem will await the development of the next generation of computers." During the 1990s, significant development of computer hardware allowed some translation of modeling to desktop computers. Forward modeling algorithms advanced too (see, e.g., Zhang et al., 1995; Zhou and Greenhalgh, 2001); however, they remained based on structured grids, which are inherently inefficient for large 3-D problems. Although the methodologies had been developed for decades, there was little attempt to move to unstructured formulations, partly because of the complexity of associated mesh generation. In part, triggered by the widespread availability of significant in-core memory on modern desktop computers, Rücker et al. (2006) exploited the emerging availability of 3-D mesh generation tools and more efficient solving methods by formulating a 3-D resistivity solution in an unstructured mesh, capable of solving the potential field problem on a mesh with several hundred thousand nodes. Their approach offers not only an accurate solution (in the main due to the ability to discretize finely near-current sources) but also a means of discretizing over a complex geometry, for example, with surface topography (e.g., Udphuay et al., 2011). Ren and Tang (2010) demonstrated how greater accuracy can be achieved by automatic mesh refinement: iteratively adapting the mesh to the resistivity distribution. The approach of Rücker et al. (2006) is arguably the current benchmark in 3-D resistivity modeling. Solutions of problems with 10^6 unknowns (voltages) appear tractable with nonspecialized computer hardware. Application to highly parallel computation is allowing solutions of even larger problems (e.g., Johnson et al., 2010).

11.08.4.1.1 IP forward modeling

The forward model for the IP problem defined in terms of complex resistivity can be computed using extensions to the DC resistivity approach, except that resistivities, voltages, and boundary conditions are expressed as complex variables. Furthermore, complex linear equation solutions are required. Kemna (2000) demonstrated the approach for the 2-D problem. Extension to 3-D is straightforward, although computer storage demands roughly double in comparison to the DC resistivity problem, which may limit application to smaller problems.

If the chargeability formulation of the IP problem is adopted, then, using the definition of apparent chargeability (see **Section 11.08.2.2.1**) from Seigel (1959), we can write (Oldenburg and Li, 1994)

$$M_a = \frac{F(\rho_{IP}) - F(\rho_0)}{F(\rho_{IP})} \qquad [30]$$

where F is the DC resistivity forward model operator, ρ_0 is the DC resistivity $\rho_{IP} = \rho_0/(1-m)$, and m is the intrinsic chargeability. Thus, from the solution of two DC resistivity forward problems, we can compute the apparent chargeability.

11.08.4.2 Inverse Modeling

Whereas the forward model computes the apparent resistivity (DC or complex) and/or apparent chargeability from a spatial distribution of resistivity (and intrinsic chargeability for the chargeability-defined IP problem), the inverse model derives the set of spatial geoelectrical properties that is consistent with the observed data (apparent resistivity and/or chargeability). As stated earlier, this was traditionally achieved for 1-D sounding data using type curves; that is, the inversion process was done by manually fitting the observed response (a sounding curve) to a series of theoretical responses. Such fitting is both subjective (since it depends on the modeler's interpretation) and labor-intensive and also impractical for 2-D and 3-D modeling. Advances in automated fitting (inverse) algorithms over the past few decades have led to the availability of a series of flexible, robust, and computationally efficient algorithms for inverse modeling of geoelectrical data for 1-D, 2-D, and 3-D problems. We highlight here some of the key developments and summarize the general approaches now widely adopted for inverse modeling of DC resistivity and IP data.

Unconstrained inverse modeling of geoelectrical data is inherently nonunique, in that there are likely to exist a large number of geoelectrical models (e.g., distributions of resistivity) that comply with the observed data. We are thus forced to constrain the inverse model, as will be discussed in this section. Furthermore, without appropriate constraints, errors (e.g., numerical rounding errors) can propagate and lead to an unstable solution.

Most geoelectrical inverse models used today are based on a least-squares fit between data and model parameters. We can express the data-model misfit as

$$\Phi_d = (\boldsymbol{d} - F(\boldsymbol{m}))^T \boldsymbol{W}_d^T \boldsymbol{W}_d (\boldsymbol{d} - F(\boldsymbol{m})) \qquad [31]$$

where $\boldsymbol{d}$ are the data (e.g., measured apparent resistivities), $F(\boldsymbol{m})$ is the set of equivalent forward model estimates with parameter set $\boldsymbol{m}$, and $\boldsymbol{W}_d$ is a data weight matrix, which, if we consider the uncorrelated measurement error case and ignore forward model errors, is a diagonal matrix with entries equal to the standard deviation of each measurement. For a 1-D DC resistivity sounding problem, $\boldsymbol{m}$ is given by a set of resistivities and associated layer thicknesses; for a 2-D resistivity imaging problem, $\boldsymbol{m}$ is a set of resistivities in a 2-D grid (or mesh).

Attempts have been made to minimize Φ_d in eqn [31] using a variety of automated curve matching procedures (e.g., Barker, 1992; Zohdy, 1989). However, many of these proved to be susceptible to slow convergence (or even divergence) of the solution. Although the damped least-squares optimization method of Marquardt (1963) can help overcome stability problems, it was Occam's method proposed by Constable et al. (1987) that offered a major breakthrough in geoelectrical inverse modeling and is fundamental to the majority of inverse solutions of DC resistivity and IP problems today. In fact, Steven Constable's original 1987 paper is currently the second most highly cited paper in the last 50 years of the journal *Geophysics*.

The method of Constable et al. (1987) searches for the smoothest model (set of parameters) that is consistent with the data. The label 'Occam's' was used by Constable et al. (1987) to emphasize the search for the simplest model (after Occam's razor). Their approach utilizes spatial regularization (Tikhonov and Arsenin, 1977) to enforce smoothing, which

also helps ensure a stable and unique solution. Constable et al. (1987) outlined their approach for 1-D problems, following which a number of smoothness-based 2-D solutions emerged (e.g., deGroot-Hedlin and Constable, 1990; Loke and Barker, 1995; Sasaki, 1989). Loke and Barker (1995) focused on efficient computation with a view to offering inversion tools that could ultimately be used in the field (at the time, this was considered to be still computationally prohibitive). Their work led to the most widely used 2-D DC resistivity inversion software (RES2DINV), which has no doubt contributed to the wide success of 2-D imaging, particularly for environmental geophysics problems. Several similar approaches emerged in the 1990s; in the succeeding text, we outline the general principle adopted by many of these.

Regularizing the minimization problem can be achieved by adding a model penalty term:

$$\Phi_{\mathrm{m}} = \boldsymbol{m}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{m} \qquad [32]$$

where $\boldsymbol{R}$ is a roughness matrix that describes a spatial connectedness of the parameter call values. For example, if we consider three parameters m_1, m_2, and m_3 in a 1-D arrangement, then if the array $\boldsymbol{R}$ is written as $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$, then Φ_{m} in eqn [32] equates to $(m_1 - m_2)^2 + (m_2 - m_3)^2$.

We then wish to seek the minimum of

$$\Phi_{\mathrm{total}} = \Phi_{\mathrm{d}} + \alpha\Phi_{\mathrm{m}} \qquad [33]$$

where α is a scalar that controls the emphasis of smoothing.

We can state a desired level of data misfit as $\Phi_{\mathrm{d}} = N$, where N is the number of measurements. In an Occam's solution, we seek to satisfy the minimization of eqn [33], subject to the largest value of α. The process is achieved by utilizing the Gauss–Newton approach, which results in the iterative solution of

$$\left(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}\boldsymbol{J} + \alpha\boldsymbol{R}\right)\triangle\boldsymbol{m} = \boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}^{\mathrm{T}}(\boldsymbol{d} - F(\boldsymbol{m}_k)) - \alpha\boldsymbol{R}\boldsymbol{m}_k \qquad [34]$$

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k + \triangle\boldsymbol{m}$$

where $\boldsymbol{J}$ is the Jacobian (or sensitivity) matrix, given by $\boldsymbol{J}_{i,j} = \partial d_i/\partial m_j$; $\boldsymbol{m}_k$ is the parameter set at iteration k; and $\triangle\boldsymbol{m}$ is the parameter update at iteration k. For the DC resistivity case, the inverse problem is typically parameterized using log-transformed resistivities.

Some implementations of the regularized solution in the preceding text have adopted a fixed scalar α; however, it is preferable to adjust this scalar through the iterative steps: initially starting with a large value and reducing the value until convergence has been reached (see, e.g., deGroot-Hedlin and Constable, 1990).

Figure 13 shows an example inversion using the approach described in the preceding text. Data from the forward model of the resistivity structure in Figure 9 (i.e., the apparent resistivities shown in Figure 10) were perturbed with 2% Gaussian noise (i.e., uncorrelated errors were added with a zero mean and a standard deviation equal to 2% of the measured apparent resistivity). These corrupted data were then inverted with a smoothness-constrained inversion. The results reveal good recovery of the three zones, at least to a depth of 6 m (three times the electrode spacing). Note that the Wenner data lead to greater lateral spreading of the vertical resistive anomaly and also a weaker magnitude of this feature. The dipole–dipole data clearly give better resolution of the boundaries (as discussed earlier), but note that a smoothness-based inversion strategy is not necessarily optimum for this model since the boundaries in the true model (Figure 9) have sharp contrasts in resistivity.

11.08.4.2.1 3-D inversion

During the 1990s, 3-D DC resistivity inversion algorithms evolved, some based on finite difference techniques (e.g., Loke and Barker, 1996a,b; Park and Van, 1991; Zhang et al., 1995)

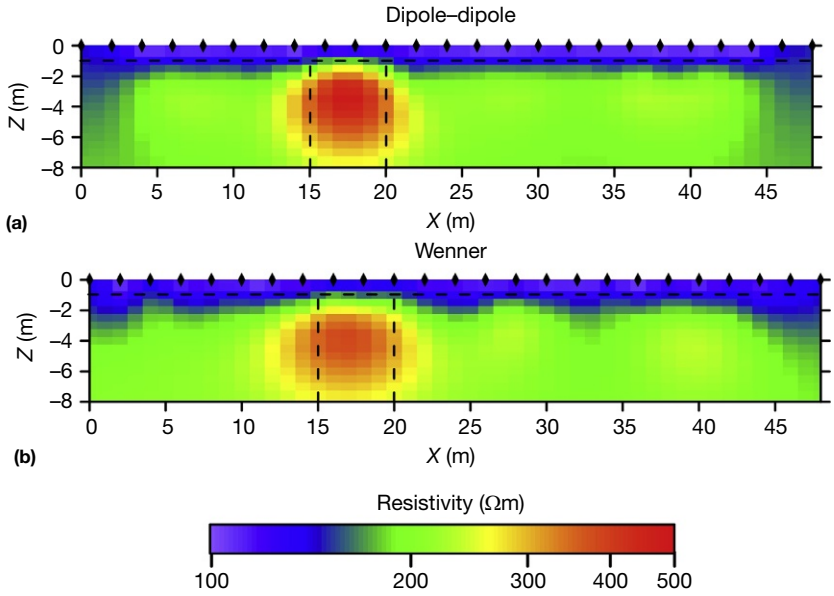


Figure 13 Inverted models for data shown in Figure 15 perturbed with 2% Gaussian noise. (a) Dipole–dipole, (b) Wenner. The dashed lines show the boundaries of the true model (as shown in Figure 9).

and others based on more flexible (but computationally more expensive) finite element solutions (e.g., Binley et al., 1996; LaBrecque et al., 1999; Sasaki, 1994). Formulation of the Jacobian matrix represents a significant computational effort, which was an early constraint for 3-D solutions, although the adjoint method allows substantial savings in computation (see, e.g., McGillivray and Oldenburg, 1990). Loke and Barker (1996b) outlined a quasi-Newton approach to make major computational savings in a 3-D inverse solution, although their results reveal some degradation of reliability in comparison with the full Gauss–Newton formulation. Storage of the entire Jacobian matrix can also be considerable for large 3-D problems, and early attempts were made to minimize such effects (e.g., Zhang et al., 1995). Solution of the full linear matrix equation in eqn [34] also precluded applications to very large grids.

Most of the early finite element-based inverse solutions used structured meshes, and thus, the key advantage over finite difference-based formulations was limited (and, in fact, they were probably less efficient computationally); however, the development of an unstructured mesh finite element solution by Günther et al. (2006) offered more accurate forward modeling (through local grid refinement) and adaptation to complex geometric problems. Coupled with this, Günther et al. (2006) exploited the reemergence of linear equation solvers that proved effective on modern computers for solving large problems. Johnson et al. (2010) had adopted similar approaches to those used by Günther et al. (2006), but by exploiting the natural parallel computational elements of the inverse problem, they have been able to apply geoelectrical inversions to very large-scale problems (e.g., 840 borehole electrodes, 150 000 measurements, and over 640 000 parameter values).

11.08.4.2.2 Time-lapse inversion

Note that the penalty function in eqn [32] could be expressed in terms of a difference relative to a reference model $\boldsymbol{m}_{\text{ref}}$:

$$\Phi_{\text{m}} = (\boldsymbol{m} - \boldsymbol{m}_{\text{ref}})^{\text{T}} \boldsymbol{R} (\boldsymbol{m} - \boldsymbol{m}_{\text{ref}}) \qquad [35]$$

Such a constraint can be considerably effective in time-lapse studied. LaBrecque and Yang (2001a) utilized this in their difference inversion scheme, which uses $\boldsymbol{m}_{\text{ref}}$ as the baseline resistivity model (prior to any change) and then adopts the following modification to eqn [34]:

$$\left(\boldsymbol{J}^{\text{T}} \boldsymbol{W}_{\text{d}}^{\text{T}} \boldsymbol{W}_{\text{d}} \boldsymbol{J} + \alpha \boldsymbol{R}\right) \triangle \boldsymbol{m} = \boldsymbol{J}^{\text{T}} \boldsymbol{W}_{\text{d}}^{\text{T}} [(\boldsymbol{d} - \boldsymbol{d}_{\text{ref}}) - (F(\boldsymbol{m}_i) - F(\boldsymbol{m}_{\text{ref}}))] - \alpha \boldsymbol{R} (\boldsymbol{m}_i - \boldsymbol{m}_{\text{ref}}) \qquad [36]$$

where $\boldsymbol{d}_{\text{ref}}$ is the baseline data vector. This approach, which effectively focuses on removing the effect of systematic errors, has proved effective in a number of time-lapse imaging studies (e.g., Doetsch et al., 2012a; LaBrecque et al., 2004). More recent attempts to formulate DC resistivity inversions in a time-lapse framework include Hayley et al. (2011), Johnson et al. (2012a), Kim et al. (2009), and Oldenborger et al. (2007a).

11.08.4.2.3 IP inversion

For inversion of IP data using the chargeability concept, model cells are parameterized in terms of DC resistivity and intrinsic chargeability, with data in the form of apparent resistivity and chargeability. Oldenburg and Li (1994) discussed three methods in which the problem can be solved. The simplest (and probably the most common in available codes) is one in which the DC resistivity problem is first solved in an iterative manner (using, e.g., the approaches described earlier), and then, the parameter chargeabilities are recovered by a linear approximation step. Since this approach requires little additional computational effort beyond that of a DC resistivity inversion, application to large 3-D problems is achievable. The major constraint with this intrinsic chargeability inverse approach is the difficulty in relating resultant values to their material properties, since intrinsic chargeability is not considered as a petrophysical property.

Alternatively, the DC resistivity inversion formulation can be translated to complex resistivity; in this case, the model may be parameterized in terms of resistivity magnitude and phase angle. Kemna (2000) outlined such a formulation for the 2-D case; Kemna et al. (2004) illustrated the approach for a series of 2-D cross borehole field studies. Kemna (2000) noted the need to add an additional stage to the inversion process that permits final refinement of the phase angle for each parameter cell, since the main stage of the inversion is dominated by the magnitude of the resistivity.

Spectral properties of the complex resistivity can be achieved by inverting each dataset for a given frequency and then applying relaxation models to each parameter cell (e.g., Kemna, 2000; Williams et al., 2009). However, it would be possible to invert spectral complex resistivity data as one single dataset (space-frequency), using regularization in both space and frequency dimensions.

Clearly, the computational demands on a complex algebra-based algorithm are greater than the DC resistivity equivalent. Computer storage is also roughly doubled as a result of the complex variables, which may limit the size of 3-D models, although Commer et al. (2011) demonstrated a formulation based on an in-phase and out-of-phase conductivity parameterization, which appears to offer a more computationally efficient solution.

11.08.4.2.4 Accounting for other a priori information in an inversion

LaBrecque and Yang (2001b) noted horizontal banding resulting from some applications of cross borehole resistivity imaging and attributed this to anisotropy in the electrical properties. Some researchers have attempted to formulate the inverse problem in terms of anisotropic resistivity (e.g., Herwanger et al., 2004; Kim et al., 2011; Zhou et al., 2009). However, it is not clear whether uniqueness of solution can be achieved with the additional parameter degrees of freedom. Furthermore, sufficient parameter discretization may remove the need to formulate in this way, since one could consider anisotropy as a different scale of heterogeneity.

Anisotropy of the spatial regularization can be easily achieved by simple modification to the roughness matrix $\boldsymbol{R}$ in eqn [32]. This does not represent anisotropy of the resistivity (as in the preceding text) but allows the user to enhance smoothing in one or more direction based on a priori information. **Figure 14** shows examples based on the inversion of data from the forward model in **Figure 9**. Here, dipole–dipole

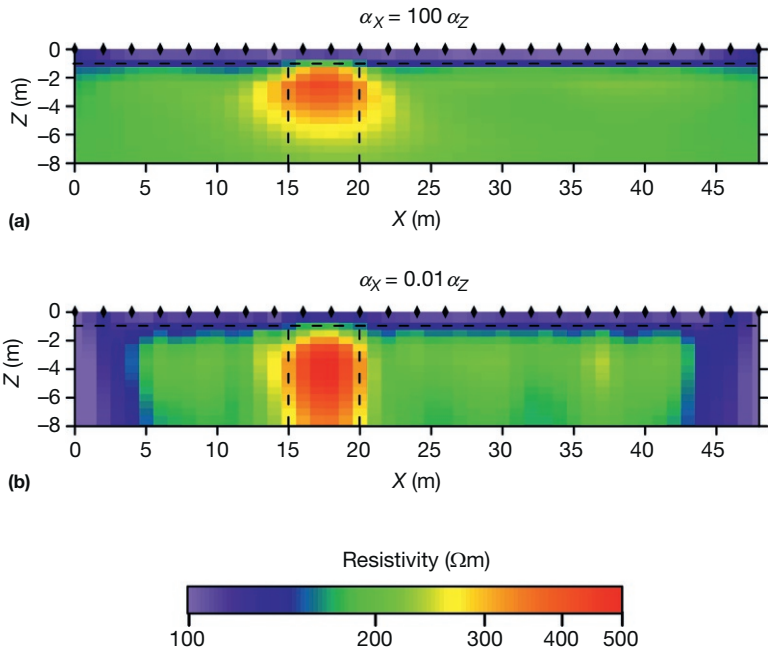


Figure 14 Anisotropic regularization. Inverted models for dipole–dipole data shown in **Figure 10(a)** perturbed with 2% Gaussian noise. (a) Enhanced horizontal smoothing. (b) Enhanced vertical smoothing. The dashed lines show the boundaries of the true model (as shown in **Figure 9**).

data were inverted with enhanced regularization in the horizontal (**Figure 14(a)**) and in the vertical (**Figure 14(b)**). These can be compared with the isotropic regularization case in **Figure 13(a)**. Note in **Figure 14(b)** that vertical conductive features appear – this is the result of the emphasis of the regularization in areas of the model where data have limited or no influence. Clearly, such effects must be treated with caution, and as discussed later, methods for appraising inverted models are essential.

Although the commonly used smoothness-constrained regularization may be applicable in a wide range of subsurface environments, there are instances where a 'blockier' structure is more appropriate (e.g., the synthetic example shown earlier where clear sharp contrasts in geoelectrical properties exist). One may consider different regularization structures (i.e., a different construction of matrix $\boldsymbol{R}$ in eqn [32]) in addition to imposing additional penalties in the inversion. de Groot-Hedlin and Constable (2004) demonstrated an approach that incorporates layer depths and lateral smoothing within the penalty function, thus enhancing sharp contrasts in resistivity. They noted, however, that their 'boundary inversion method' demonstrated occasional instabilities and was prone to being trapped in local minima. Blaschek et al. (2008) used, in their analysis of surface IP data, the minimum gradient support approach of Portniaguine and Zhdanov (1999), which penalizes parameter differences between adjacent cells, thus resulting in sharp contrasting geoelectrical images.

The spatial regularization can also be easily modified to account for known discontinuities by removing smoothing at such locations. Slater and Binley (2006) took this more simplistic approach in their study of geoelectrical imaging of permeable reactive barriers, where, in this case, the engineered structure boundaries were known. Doetsch et al. (2012b) adopted a similar strategy, in this case using ground-penetrating radar data to provide a priori information about lithologic boundaries, that we assumed to have electrical property contrasts. Such simplistic approaches offer the benefits of well-behaved (in terms of convergence and global minima) smoothness-constrained inversions (within the regularization disconnected regions) while also retaining sharp contrasts in known areas. Alternatively, one may adopt image processing techniques to postprocess smoothness-constrained inversions (e.g., Nguyen et al., 2005) or use such image processing techniques to reveal regions where regularization disconnection could be considered. Clearly, inappropriate designation of disconnected regions is likely to lead to erroneous inverse model results and thus must be carefully applied.

Figure 15 shows the effect of applying regularization disconnects to the synthetic model used earlier. In this case, we invert the dipole–dipole data using the same settings as in **Figure 13(a)** except for a minor modification to the roughness matrix that removes smoothing along the boundary of the 500 Ωm feature in the original model (see **Figure 9**). Thus, smoothing still exists across parameter cells within this zone and outside the zone. The effect is striking, demonstrating how effective such a priori information can be.

More sophisticated uses of a priori knowledge are emerging in geoelectrical inversions that account for spatial constraints in terms of geostatistical functions, rather than a simplistic smoothing function. Linde et al. (2006) demonstrated such an approach in their joint inversion of cross borehole resistivity and radar data, using electromagnetic induction well logs from a number of boreholes to formulate a geostatistical model. Johnson et al. (2012b) followed a similar philosophy, in their case by also constraining resistivity values at boreholes, based on well-logged values.

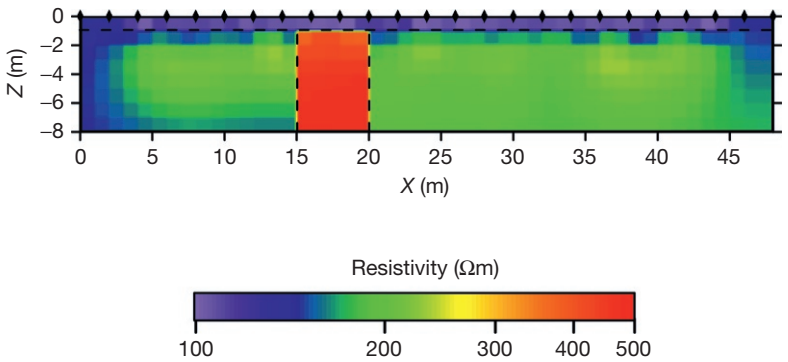


Figure 15 Effect of regularization disconnects. Inverted models for dipole–dipole data shown in **Figure 10(a)** perturbed with 2% Gaussian noise. At boundaries of the high resistive zone (see **Figure 9**), a disconnect in regularization is applied. All other factors in the inversion are the same as **Figure 13(a)**. The dashed lines show the boundaries of the true model (as shown in **Figure 9**).

The smoothness (or Occam's) inversion remains the most widely used inversion approach in geoelectrical modeling, not only because of its good convergence and stability but also because the models derived from such inversion can be visually appealing. It is important to remember, however, that the smoothness constraint is a constraint and implies a priori knowledge that this is appropriate. Even though the inverted model can be considered unique, this is only subject to the constraints applied. It, therefore, could be considered as one of an infinite number of geoelectrical models.

11.08.4.2.5 Image appraisal

Since the sensitivity of each measurement is a complex spatial function (see, e.g., **Figure 2**), the spatial resolution of an inversion of geoelectrical data should be assessed prior to any interpretation of results. In most 2-D surface imaging applications, a trapezium boundary is constructed following a similar pattern to that shown in the pseudosection (e.g., **Figure 10**); however, variation in resistivity will impact on the spatial sensitivity and such linear boundaries may be inappropriate. Furthermore, for more complex geometric arrangements (e.g., borehole-based imaging or 3-D imaging), such simplistic approaches are inadequate.

One approach that is widely appreciated in general inverse theory (e.g., Menke, 1984) is the resolution matrix, $\mathcal{R}$, which is defined by (e.g., Day-Lewis et al., 2005)

$$\boldsymbol{m} = \mathcal{R}\boldsymbol{m}_{\text{true}} \qquad [37]$$

where $\boldsymbol{m}$ is the inverted parameter set and $\boldsymbol{m}_{\text{true}}$ is the (unknown) true parameter set. Thus, the ideal structure of $\mathcal{R}$ is the identity matrix, since this implies a perfect mapping of true and inverted parameter vectors. Any deviation from the identity matrix reveals the lack of sensitivity of the parameter values to the measured data, coupled with the effect of smoothing and other regularization.

The resolution matrix can be computed, for the formulation in eqn [34], using

$$\left(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}\boldsymbol{J} + \alpha\boldsymbol{R}\right)\mathcal{R} = \boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}\boldsymbol{J} \qquad [38]$$

where the Jacobian, $\boldsymbol{J}$, has been computed based on the final (inverted) parameter set and the regularization scalar, α, is the value at the end of the inversion.

The computation of $\mathcal{R}$ requires significant effort: the formation and solution of M sets of equations, each of size $M \times M$, where M is the number of parameters. Consequently, $\mathcal{R}$ is rarely presented in applications, despite being a very insightful matrix (e.g., Ramirez et al., 1993).

Because of the computation burden of computing $\mathcal{R}$, alternative approaches have been adopted in some studies. Park and Van (1991) and Kemna (2000), for example, had used a sensitivity map, $\boldsymbol{S}$, expressed, for parameter j, as

$$S_j = \left(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}\boldsymbol{J}\right)_{jj} \qquad [39]$$

Oldenburg and Li (1999) developed a different and effective approach to establish a depth of investigation based on the inversion of a dataset using two starting models $\boldsymbol{m}_{\text{ref}}^{(\text{a})}$ and $\boldsymbol{m}_{\text{ref}}^{(\text{b})}$ (Oldenburg and Li (1999) recommend that these vary by a factor of 5–10) to which a penalty function is applied in the inversion, in such a way that the inverted parameters, $\boldsymbol{m}^{(\text{a})}$ and $\boldsymbol{m}^{(\text{b})}$, respectively, are constrained to the starting model unless driven by sensitivity to the data. The depth of investigation (DOI) is then computed for cell j as

$$\text{DOI}_j = \frac{\boldsymbol{m}_j^{(\text{a})} - \boldsymbol{m}_j^{(\text{b})}}{\boldsymbol{m}_{\text{ref}\,j}^{(\text{a})} - \boldsymbol{m}_{\text{ref}\,j}^{(\text{b})}} \qquad [40]$$

In order to implement this approach, we require an additional penalty term to constrain the model to the reference parameter vector. This can be achieved by modifying eqn [34] to

$$\left(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}\boldsymbol{J} + \alpha\boldsymbol{R} + \alpha_{\mathrm{s}}\boldsymbol{I}\right)\triangle\boldsymbol{m} = \boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}_{\mathrm{d}}^{\mathrm{T}}(\boldsymbol{d} - F(\boldsymbol{m}_i)) - \alpha\boldsymbol{R}\boldsymbol{m}_i - \alpha_{\mathrm{s}}\mathbf{m}_{\text{ref}} \qquad [41]$$

where α_{s} is an additional penalty scalar and $\boldsymbol{I}$ is the identity matrix.

In this approach, a small value of DOI indicates a high degree of sensitivity to the measurements. Oldenburg and Li (1999) suggested that a value of the depth of investigation equal to 0.1–0.2 represents a reasonable upper threshold for satisfactory sensitivity (an equivalent threshold is also required for the resolution matrix approach). Note that the DOI formulation can be applied to resistivity and IP inversions and is easily extended to 3-D problems, as demonstrated by Oldenborger et al. (2007b).

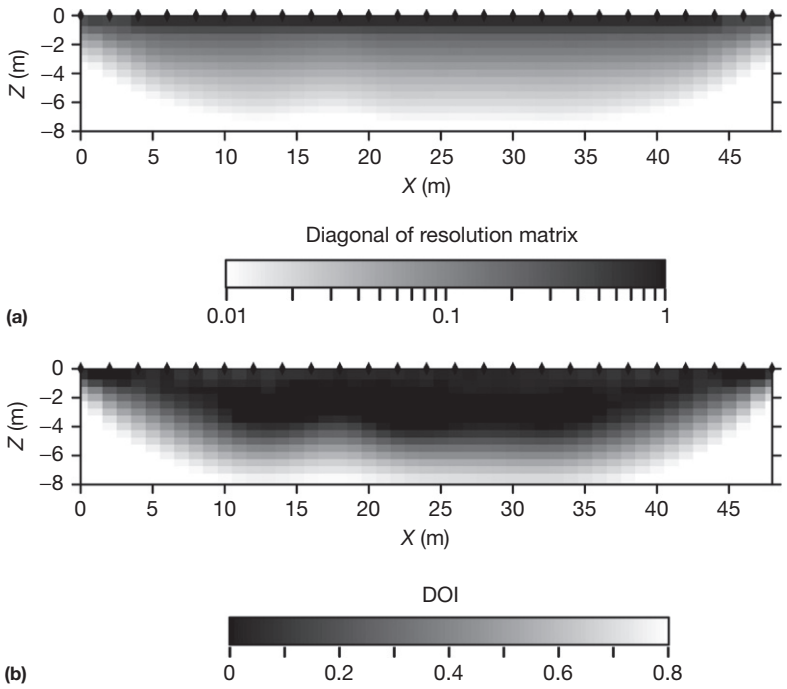


Figure 16 Two example methods for inverse model appraisal. (a) Diagonal of resolution matrix. (b) Oldenburg and Li's depth of investigation (DOI). Results based on dipole–dipole dataset for forward model in **Figure 9**.

The resolution matrix and depth of investigation approaches are illustrated in **Figure 16** for the dipole–dipole dataset from the model in **Figure 9**. For the depth of investigation, two inverse models were computed with a reference resistivity of 100 and 200 Ωm, using the formulation in eqn [41] with $\alpha_s = 10\alpha$.

11.08.4.2.6 Data and model errors

In eqn [31], we included the matrix W_d to allow data to be weighted according to their reliability. As stated earlier, under the assumption of uncorrelated data errors, this can be assumed to be a diagonal matrix with each entry equal to the reciprocal of the standard deviation of a measurement. This also assumes that the model is error-free. Applying such weights not only permits the differential weighting of poor/good data but also allows the objective definition of a satisfactory convergence: $\Phi_d = N$, where N is the number of measurements. However, in numerous applications of geoelectrical inversion, such a criterion is not adopted, and often, the user displays an equivalent uniform data error that the final model represents.

Despite its significance, few studies have explored the impact of errors on geoelectrical inversions. Binley et al. (1995) and LaBrecque et al. (1996b) are rare examples of such studies, both highlighting the significance of incorrect data error estimation on the final inverted model. **Figure 17** shows how critical the noise level estimates are. Here, we perturbed the dipole–dipole dataset (**Figure 10(a)**) with 5% uncorrelated Gaussian noise and then inverted with an assumed noise level. In **Figure 17(b)**, the correct noise level is assumed and the model is reasonably recovered, albeit with slightly less contrast in resistivity than the 2% noise case in **Figure 13(a)** (as expected). If we assume a worse noise level (in this case 10%) than the true level, then the model is poor (**Figure 17(a)**): not all the information in the data is recovered by the inversion. If, however, we underestimate the noise level and assume a lower value (in this case 2%), then the final model reveals high variability (**Figure 17(c)**) due to the 'overfitting.' These examples demonstrate the importance of (a) assessing error levels and (b) accounting for such error levels in the inversion process. Failure to do so could result in either (i) failure to exploit all the information in the data or (ii) incorrect interpretation of the subsurface geoelectrical structure.

Morelli and LaBrecque (1996) developed an adaptive data weighting scheme that allows the inversion process to adjust weights throughout the inversion process. Such schemes are highly effective in guaranteeing convergence, since measurements that appear as outliers in the inversion become weighted less. These methods have, no doubt, helped immensely in producing extremely robust inversion tools, but they must be used with care. Data outliers may exist because the model is incorrect (e.g., the assumption of 2-D resistivity or incorrectly assuming inclination of boreholes in a borehole imaging survey): rejection or reduced weight of such data may then lead to a false representation of the subsurface.

Finally, it is worth noting that as instruments and operational procedures (e.g., the appropriate assessment of measurement errors) improve, data quality is enhanced. However, constraints in the ability to accurately model the geoelectrical response of the subsurface still exist (although clearly, they are less significant than in earlier applications). Consequently, we may find that errors in the forward model need to be recognized more and appropriately accounted for in the data misfit

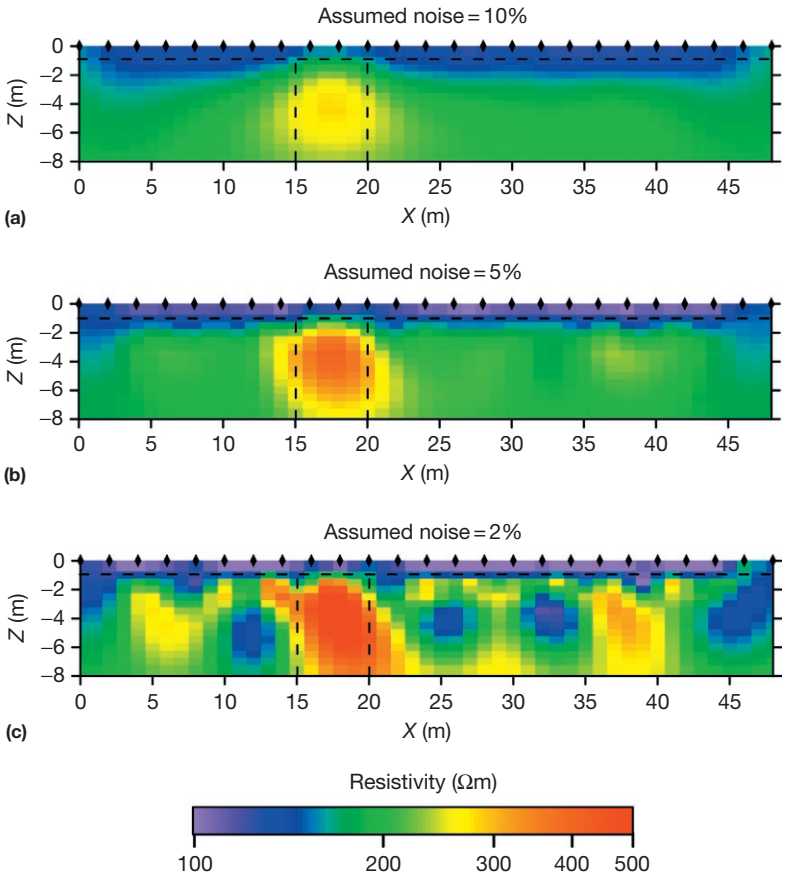


Figure 17 Effect of noise on inverted model. Inversion of dipole–dipole dataset from the model in **Figure 9** with 5% noise added. The model was derived through inversion with different assumed noise levels. (a) 10%. (b) 5% (i.e., correct level). (c) 2%.

objective function. This is particularly the case for 3-D problems, where computational constraints still exist. Assessing forward models for half-space problems with no topography is trivial since analytic solutions for homogenous problems are easy to compute. For other problems, one is forced to consider mesh refinement as a means of computing a reference 'accurate' solution for error computation.

11.08.5 Summary and Outlook

Electrical methods have evolved in geophysics since their inception over a hundred years ago, although the measurement principle in modern applications is very similar to that originally adopted. The basic four-electrode measurement can be carried out in a wide range of configurations on the ground surface, or at depth, allowing relatively rapid profiling or imaging of the subsurface. Recent advances in computer-controlled measurement devices, coupled with the development of robust data inversion algorithms and wide availability of suitable computer processing capability, have led to the almost routine application of 2-D resistivity imaging, particular for near-surface environmental and engineering problems. The development of multichannel instrumentation, allowing concurrent voltage measurements on an array of electrodes, has led to more applications of 3-D imaging surveys. 3-D imaging algorithms have been available for some time, but their application has been limited due to the computational constraints. The recent rapid growth in the availability of relatively large computer memory on personal computers, coupled with processor performance, has meant that 3-D imaging is now achievable for problems involving several hundred electrodes, without the need for specialized computer hardware. Continued exploitation of parallel computing architecture will no doubt help in expanding the size of problems that 3-D electrical imaging can be applied to. Data processing algorithms have also emerged to include approaches for dealing with time-lapse studies, effectively allowing 4-D imaging.

IP, although established for some time, has received a recent growth in popularity, particularly for environmental geophysics applications. There still remains immense uncertainty in some of the subsurface processes that this method is sensitive to, but there is growing recognition that for some (not all) applications, IP can offer additional information beyond a DC resistivity survey.

Despite these advances, challenges still remain in the application of electrical methods:

- Data collection time can still be a constraint for large 3-D arrays. Most measurement systems are still based on mechanical switches, and many new multichannel

instruments do not have full array switching capability. Both factors constrain the data collection speed, which can be problematic for studying time-lapse processes.

- The majority of applications of electrical methods do not appear to involve any assessment of uncertainty in the derived image of the subsurface. Most adopt a single inversion approach (e.g., regularization scheme) and compute a single result, with no apparent appreciation that this represents the 'best' (but not only) model that is consistent with the data, subject to the constraints of the imposed regularization.
- Data errors and (forward) model errors are rarely reported, and their impacts on the data inversion are not universally assessed, yet the impact of inappropriate error assumptions can be critical. Repeatability and reciprocity checks may not necessarily account for all measurement errors, and as instruments and measurement practice improve, modeling errors may dominate the uncertainty in the final model.
- Image appraisal approaches have been developed, and these can offer great insight into the reliability of the final model. Again, these approaches are not routinely used, and when they are, the selection of appropriate thresholds makes their use somewhat subjective. Furthermore, application of image appraisal schemes to IP studies are lacking: we need additional insight into the value of information from IP surveys, and image appraisal is one element of this.

One hundred years on from the inception of electrical methods in geophysics, we now have the tools in order to (at least in theory) be able to image the subsurface in 5-D (3-D space, time-lapse, and frequency dependency). Electrical methods have been established as extremely powerful and versatile and applicable to a wide range of problems, particularly for near-surface applications. Their use will no doubt develop, and aided perhaps with further exchange across other disciplines (e.g., biomedicine), algorithm and instrumentation will continue to advance in order to tackle some of the current constraints, including those listed in the preceding text.

Acknowledgments

I am grateful to Torleif Dahlin for bringing to my attention the work of the two British pioneers Alfred Williams and Leo Daft and to Robert Vernon for providing me with historical information about the work of Williams and Daft. I also acknowledge Judy Robinson, Neil Terry, and Markus Wehrer for their reviewers of an earlier version of the chapter.

References

Allen D and Merrick N (2007) Robust 1D inversion of large towed geo-electric array datasets used for hydrogeological studies. *Exploration Geophysics* 38: 50–59.

Anderson WL (1979) Numerical integration of related Hankel transforms of orders 0 and 1 by adaptive digital filtering. *Geophysics* 44: 1287–1305.

André F, van Leeuwen C, Saussez S, et al. (2012) High-resolution imaging of a vineyard in south of France using ground-penetrating radar, electromagnetic induction and electrical resistivity tomography. *Journal of Applied Geophysics* 78: 113–122.

Barber DC and Brown BH (1984) Applied potential tomography. *Journal of Physics E: Scientific Instruments* 17(9): 723–733.

Barker RD (1979) Signal contributions and their use in resistivity studies. *Geophysical Journal of the Royal Astronomical Society* 59(1): 123–129.

Barker RD (1992) A simple algorithm for electrical imaging of the subsurface. *First Break* 10(2): 53–62.

Barsoukov E and Macdonald JR (2005) *Impedance Spectroscopy. Theory, Experiment and Applications*. Wiley, 595 pp.

Beasley C and Ward S (1986) Three-dimensional mise-à-la-masse modeling applied to mapping fracture zones. *Geophysics* 51(1): 98–113.

Bevc D and Morrison HF (1991) Borehole-to-surface electrical resistivity monitoring of a salt water injection experiment. *Geophysics* 56(6): 769–777.

Bhattacharya B, Gupta D, Banerjee B, and Shalivahan (2001) Mise-a-la-masse survey for an auriferous sulfide deposit. *Geophysics* 66(1): 70–77.

Bing Z and Greenhalgh SA (2000) Cross-hole resistivity tomography using different electrode configurations. *Geophysical Prospecting* 48: 887–912.

Binley A, Cassiani G, Middleton R, and Winship P (2002a) Vadose zone model parameterisation using cross-borehole radar and resistivity imaging. *Journal of Hydrology* 267(3–4): 147–159.

Binley A, Pinheiro P, and Dickin F (1996) Finite element based three-dimensional forward and inverse solvers for electrical impedance tomography. In: *Proceedings of Colloquium on Advances in Electrical Tomography, Computing and Control Division, IEE, Digest No. 96/143, Manchester, June, 1996*, pp. 6/1–6/3.

Binley AM, Ramirez A, and Daily W (1995) Regularised image reconstruction of noisy electrical resistance tomography data. In: Beck MS, et al. (eds.) *Process Tomography, Proceedings of Fourth Workshop of the European Concerted Action on Process Tomography, Bergen, Norway*, pp. 401–410.

Binley A, Winship P, West LJ, Pokar M, and Middleton R (2002b) Seasonal variation of moisture content in unsaturated sandstone inferred from borehole radar and resistivity profiles. *Journal of Hydrology* 267(3–4): 160–172.

Blaschek R, Hördt A, and Kemna A (2008) A new sensitivity-controlled focusing regularization scheme for the inversion of induced polarization data based on the minimum gradient support. *Geophysics* 73(2): F45–F54.

Bleil D (1953) Induced polarization: A method for geophysical prospecting. *Geophysics* 18(3): 636–662.

Blome M (2010) *Efficient Measurement and Data Inversion Strategies for Large Scale Geoelectric Surveys*. PhD Thesis, ETH Zurich.

Blome M, Maurer H, and Greenhalgh S (2011) Geoelectric experimental design – Efficient acquisition and exploitation of complete pole-bipole data sets. *Geophysics* 76(1): F15–F26.

Bodmer K, Ward SH, and Morrison HF (1968) On induced electrical polarization and groundwater. *Geophysics* 33: 805–821.

Börner FD, Gruhne M, and Schön JH (1993) Contamination indications derived from electrical properties in the low frequency range. *Geophysical Prospecting* 41: 83–98.

Börner FD and Schön JH (1991) A relation between the quadrature component of electrical conductivity and the specific surface area of sedimentary rocks. *The Log Analyst* 32: 612–613.

Bradbury KR and Taylor RW (1984) Determination of the hydrogeological properties of lakebeds using offshore geophysical surveys. *Ground Water* 22: 690–695.

Cassiani G, Bruno V, Villa A, Fusi N, and Binley AM (2006) A saline trace test monitored via time-lapse surface electrical resistivity tomography. *Journal of Applied Geophysics* 59: 244–259.

Christensen NB and Sørensen K (2001) Pulled array continuous electrical sounding with an additional inductive source: An experimental design study. *Geophysical Prospecting* 49: 241–254.

Clark A (1990) *Seeing Beneath the Soil. Prospecting Methods in Archaeology*. London: Routledge.

Coggon J (1971) Electromagnetic and electrical modeling by the finite element method. *Geophysics* 36(1): 132–155.

Cole KS and Cole RH (1941) Dispersion and absorption in dielectrics. *Journal of Chemical Physics* 9: 341–351.

Commer M, Newman G, Williams K, and Hubbard S (2011) 3D induced-polarization data inversion for complex resistivity. *Geophysics* 76(3): F157–F171.

Constable S, Parker R, and Constable C (1987) Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics* 52(3): 289–300.

Dahlin T (1996) 2D resistivity surveying for environmental and engineering applications. *First Break* 14(7): 275–283.

Dahlin T (2000) Short note on electrode charge-up effects in DC resistivity data acquisition using multi-electrode arrays. *Geophysical Prospecting* 48: 181–187.

Dahlin T (2001) The development of DC resistivity imaging techniques. *Computers and Geosciences* 27: 1019–1029.

Dahlin T and Leroux V (2012) Improvement in time-domain induced polarization data quality with multi-electrode systems by separating current and potential cables. *Near Surface Geophysics* 10(6): 545–656.

Dahlin T, Leroux V, and Nissen J (2002) Measuring techniques in induced polarisation imaging. *Journal of Applied Geophysics* 50: 279–298.
Dahlin T and Loke MH (1997) Quasi-3D resistivity imaging/mapping of three dimensional structures using two dimensional DC resistivity techniques. In: *Proceedings of the 3rd Meeting of the Environmental and Engineering Geophysical Society, Aarhus, Denmark,* pp. 143–146.
Dahlin T and Zhou B (2004) A numerical comparison of 2D resistivity imaging with 10 electrode arrays. *Geophysical Prospecting* 52: 379–398.
Daily W and Owen E (1991) Cross borehole resistivity tomography. *Geophysics* 56: 1228–1235.
Daily W, Ramirez A, and Binley A (2004a) Remote monitoring of leaks in storage tanks using electrical resistance tomography: Application at the Hanford site. *Journal of Environmental and Engineering Geophysics* 9(1): 11–24.
Daily W, Ramirez A, Binley A, and LaBrecque D (2004b) Electrical resistance tomography. *The Leading Edge* 23(5): 438–442.
Daily W, Ramirez A, LaBrecque D, and Nitao J (1992) Electrical resistivity tomography of vadose water movement. *Water Resources Research* 28(5): 1429–1442.
Daniels J (1977) Three-dimensional resistivity and induced polarization modeling using buried electrodes. *Geophysics* 42(5): 1006–1019.
Day-Lewis F, Singha K, and Binley A (2005) On the limitations of applying petrophysical models to tomograms: A comparison of correlation loss for cross-hole electrical-resistivity and radar tomography. *Journal of Geophysical Research* 110(B8): B08206.
DeGroot-Hedlin C and Constable S (1990) Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics* 55(12): 1613–1624.
deGroot-Hedlin C and Constable S (2004) Inversion of magnetotelluric data for 2D structure with sharp resistivity contrasts. *Geophysics* 69(1): 78–86.
Dey A and Morrison HF (1979) Resistivity modeling for arbitrarily shaped three-dimensional structures. *Geophysics* 44: 753–780.
Dias CA (2000) Developments in a model to describe low-frequency electrical polarization of rocks. *Geophysics* 65(2): 437–451.
Dieter K, Paterson N, and Grant F (1969) IP and resistivity type curves for three-dimensional bodies. *Geophysics* 34(4): 615–632.
Doetsch J, Coscia I, Greenhalgh S, Linde N, Green A, and Günther T (2010) The borehole-fluid effect in electrical resistivity imaging. *Geophysics* 75(4): F107–F114.
Doetsch J, Linde N, Pessognelli M, Green AG, and Günther T (2012a) Constraining 3D electrical resistance tomography with GPR data for improved aquifer characterization. *Journal of Applied Geophysics* 78: 68–76.
Doetsch J, Linde N, Vogt T, Binley A, and Green A (2012b) Imaging and quantifying salt tracer transport in a riparian groundwater system by means of 3D ERT monitoring. *Geophysics* 77(5): B207–B218.
Edwards LS (1977) A modified pseudosection for resistivity and IP. *Geophysics* 42(5): 1020–1036.
Fiandaca G, Auken E, Christiansen AV, and Gazoty A (2012) Time-domain-induced polarization: Full-decay forward modeling and 1D laterally constrained inversion of Cole–Cole parameters. *Geophysics* 77(3): E213–E225.
Flores Orozco A, Williams KH, Long PE, Hubbard SS, and Kemna A (2011) Using complex resistivity imaging to infer biogeochemical processes associated with bioremediation of an uranium-contaminated aquifer. *Journal of Geophysical Research* 116: G03001.
Furman A, Ferré T, and Warrick A (2004) Optimization of ERT surveys for monitoring transient hydrological events using perturbation sensitivity and genetic algorithms. *Vadose Zone Journal* 3: 1230–1239.
Gafney C and Gater J (2003) *Revealing the Buried Past. Geophysics for Archaeologists*. Stroud, UK: Tempus Publishing.
Gazoty A, Fiandaca G, Pedersen J, Auken E, and Christiansen AV (2013) Data repeatability and acquisition techniques for time-domain spectral induced polarization. *Near Surface Geophysics* 11. http://dx.doi.org/10.3997/1873-0604.2013013.
Griffiths DH and Turnbull J (1985) A multi-electrode array for resistivity surveying. *First Break* 3: 16–20.
Griffiths DH, Turnbull J, and Olayinka AI (1990) Two-dimensional resistivity mapping with a computer controlled array. *First Break* 8: 121–129.
Günther T, Rücker C, and Spitzer K (2006) Three-dimensional modelling and inversion of dc resistivity data incorporating topography – II. Inversion. *Geophysical Journal International* 166: 506–517.
Habberjam GM (1972) The effects of anisotropy on square array resistivity measurements. *Geophysical Prospecting* 20: 249–266.
Hayley K, Pidlisecky A, and Bentley LR (2011) Simultaneous time-lapse electrical resistivity inversion. *Journal of Applied Geophysics* 75: 401–411.
Herwanger JV, Pain CC, Binley A, de Oliveira CRE, and Worthington MH (2004) Anisotropy resistivity tomography. *Geophysical Journal International* 158: 409–425.
Hohmann GW (1988) Numerical modeling for electromagnetic methods of geophysics. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics. Theory*, vol. 1, pp. 313–363. Tulsa, OK: Society of Exploration Geophysicists.
Johnson IM (1984) Spectral induced polarization as determined through time-domain measurements. *Geophysics* 49: 1993–2003.
Johnson TC, Slater LD, Ntarlagiannis D, Day-Lewis FD, and Elwaseif M (2012a) Monitoring groundwater-surface water interaction using time-series and time-frequency analysis of transient three-dimensional electrical resistivity changes. *Water Resources Research* 48: W07506.
Johnson T, Versteeg R, Rockhold M, et al. (2012b) Characterization of a contaminated wellfield using 3D electrical resistivity tomography implemented with geostatistical, discontinuous boundary, and known conductivity constraints. *Geophysics* 77(6): EN85–EN96.
Johnson TC, Versteeg RJ, Ward A, Day-Lewis FD, and Revil A (2010) Improved hydrogeophysical characterization and monitoring through parallel modeling and inversion of time-domain resistivity and induced-polarization data. *Geophysics* 75(4): WA27–WA41.
Jones DP (2002) *Investigations of Clay-Organic Reactions Using Complex Resistivity*. M.S. Thesis, USA: Colorado School of Mines, 403 pp.
Keery J, Binley A, Elshenawy A, and Clifford J (2012) Markov chain Monte Carlo estimation of distributed Debye relaxations in spectral induced polarization. *Geophysics* 77(2): E159–E170.
Kemna A (2000) Tomographic Inversion of Complex Resistivity. PhD Thesis, *Theory and Application*. Germany: Bochum University, 176 pp.
Kemna A, Binley A, and Slater L (2004) Crosshole IP imaging for engineering and environmental applications. *Geophysics* 69(1): 97–107.
Kemna A, Rakers E, and Binley A (1997) Application of complex resistivity tomography to field data from a kerosene-contaminated site. In: *Proceedings of 3rd Meeting of the Environmental and Engineering Geophysics Society (European Section),* 151–154.
Keys WS (1990) Borehole geophysics applied to ground-water investigations: U.S. Geological Survey Techniques of Water-Resources Investigation, Book 2, Chapter E2, 150 pp.
Kim J-H, Yi M-J, Cho S-J, Son J-S, and Song W-K (2011) Anisotropic crosshole resistivity tomography for ground safety analysis of a high-storied building over an abandoned mine. *Journal of Environmental and Engineering Geophysics* 11(4): 225–235.
Kim J-H, Yi MJ, Park S-G, and Kim JG (2009) 4-D inversion of DC resistivity monitoring data acquired over a dynamically changing earth model. *Journal of Applied Geophysics* 68: 522–532.
Koefoed O (1979) *Geosounding Principles, Vol. 1, Resistivity Sounding Measurements*. Amsterdam: Elsevier Science.
Komorov VA (1972) *The Electrical Prospecting with Induced Polarization Method*. Leningrad: Nedra, 391 pp.
Kruschwitz S and Yaramanci U (2004) Detection and characterisation of the disturbed rock zone in claystone with the complex resistivity method. *Journal of Applied Geophysics* 57: 63–79.
LaBrecque D and Daily W (2008) Assessment of measurement errors for galvanic-resistivity electrodes of different composition. *Geophysics* 73(2): F55–F64.
LaBrecque D, Heath G, Sharpe R, and Versteeg R (2004) Autonomous monitoring of fluid movement using electrical resistivity tomography. *Journal of Environmental and Engineering Geophysics* 9: 167–176.
LaBrecque DJ, Miletto M, Daily W, Ramirez A, and Owen E (1996a) The effects of noise on Occam's inversion of resistivity tomography data. *Geophysics* 61: 538–548.
LaBrecque DJ, Morelli G, Daily W, Ramirez A, and Lundegard P (1999) Occam's inversion of 3D ERT data. In: Spies B (ed.) *Three-Dimensional Electromagnetics*, pp. 575–590. Tulsa, OK: Society of Exploration Geophysicists.
LaBrecque DJ, Ramirez AL, Daily WD, Binley AM, and Schima SA (1996b) ERT monitoring of environmental remediation processes. *Measurement Science and Technology* 7: 375–383.
LaBrecque DJ and Ward SH (1990) Two-dimensional cross-borehole resistivity model fitting. In: Ward SH (ed.) *Geotechnical and Environmental Geophysics*, vol. III, pp. 51–74. Tulsa, OK: Society of Exploration Geophysicists.
LaBrecque D and Yang X (2001a) Difference inversion of ERT data: Fast inversion method for 3D in situ monitoring. *Journal of Environmental and Engineering Geophysics* 6(2): 83–89.
LaBrecque D and Yang X (2001b) The effects of anisotropy on ERT images for Vadose zone monitoring. *Symposium on the Application of Geophysics to Engineering and Environmental Problems*, VZC2–VZC2.
Lane JW Jr., Haeni FP, and Watson WM (1995) Use of a square-array direct-current resistivity method to detect fractures in crystalline bedrock in New Hampshire. *Ground Water* 33(3): 476–485.

Lesmes DP and Friedman SP (2005) Relationships between the electrical and hydrogeological properties of rocks and soils. In: Rubin Y and Hubbard SS (eds.) *Hydrogeophysics*, pp. 87–128. Dordrecht, The Netherlands: Springer, Chapter 4.

Linde N, Binley A, Tryggvason A, Pedersen L, and Revil A (2006) Improved hydrogeophysical characterization using joint inversion of crosshole electrical resistance and ground penetrating radar traveltime data. *Water Resources Research* 42(12): W04410.

Loke MH and Barker RD (1995) Least-squares deconvolution of apparent resistivity pseudosections. *Geophysics* 60(6): 1682.

Loke MH and Barker RD (1996a) Practical techniques for 3D resistivity surveys and data inversion. *Geophysical Prospecting* 44: 499–523.

Loke MH and Barker RD (1996b) Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method. *Geophysical Prospecting* 44: 131–152.

Loke MH, Wilkinson PB, and Chambers JE (2010) Parallel computation of optimized arrays for 2D electrical imaging surveys. *Geophysical Journal International* 183: 1302–1315.

Lowry T, Allen MB, and Shive PN (1989) Singularity removal: A refinement of resistivity modeling techniques. *Geophysics* 54: 766–774.

Lytle RJ and Dines KA (1978) An impedance camera: A system for determining the spatial variation of electrical conductivity. *Lawrence Livermore National Laboratory Report UCRL-52413*, Livermore, USA.

Lytle RJ and Dines DK (1980) Iterative ray tracing between boreholes for underground image reconstruction. *IEEE Transactions on Geoscience and Remote Sensing* GE-18: 234–239.

Major J and Silic J (1981) Restrictions on the use of Cole–Cole dispersion models in complex resistivity interpretation. *Geophysics* 41: 916–931.

Mansinha L and Mwenifumbo C (1983) A mise-à-la-masse study of the Cavendish Geophysical Test Site. *Geophysics* 48(9): 1252–1257.

Marescot L, Palma Lopes S, Lagabrielle R, and Chapellier D (2002) Designing surface-to-borehole electrical resistivity tomography surveys using the Frechet derivative. In: *Proceedings of 8th Meeting of the Environmental and Engineering Geophysical Society – European Section,* pp. 289–292.

Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* 11: 431–441.

Martinho E and Almeida F (2006) 3D behavior of contamination in landfill sites using 2D resistivity/IP imaging: Cases studies in Portugal. *Environmental Geology* 49: 1071–1078.

McGillivray PR and Oldenburg DW (1990) Methods for calculating Frechet derivatives and sensitivities for the non-linear inverse problem: A comparative study. *Geophysical Prospecting* 38: 499–524.

Menke W (1984) *Geophysical Data analysis: Discrete Inverse Theory*. New York: Academic Press.

Morelli G and LaBrecque DJ (1996) Advances in ERT modeling. *European Journal of Environmental and Engineering Geophysics* 1: 171–186.

Morgan FD and Lesmes DP (1994) Inversion for dielectric-relaxation spectra. *Journal of Chemical Physics* 100: 671–681.

Mwakanyamale K, Slater L, Binley A, and Ntarlagiannis D (2012) Lithologic imaging using induced polarization: Lessons learned from the Hanford 300 area. *Geophysics* 77: 397–409.

Nabighian MN and Elliot CL (1976) Negative induced-polarization effects from layered media. *Geophysics* 41(6A): 1236–1255.

Nguyen F, Garambois S, Jongmans D, Pirard E, and Loke MH (2005) Image processing of 2D resistivity data for imaging faults. *Journal of Applied Geophysics* 57: 260–277.

Nimmer RE and Osiensky JL (2001) Direct current and self-potential monitoring of an evolving plume in partially saturated fractured rock. *Journal of Hydrology* 267: 258–272.

Nimmer RE, Osiensky JL, Binley AM, and Williams BC (2008) Three-dimensional effects causing artifacts in two-dimensional, cross-borehole, electrical imaging. *Journal of Hydrology* 359: 59–70.

Nordsiek S and Weller A (2008) A new approach to fitting induced-polarization spectra. *Geophysics* 73: F235–F245.

Nyquist JE and Roth MJS (2005) Improved 3D pole-dipole resistivity surveys using radial measurement pairs. *Geophysical Research Letters* 32: L21416.

Ogilvy RD, Kuras O, Meldrum PI, et al. (2009) Automated time-lapse electrical resistivity tomography (ALERT) for monitoring coastal aquifers. *Near Surface Geophysics* 7: 367–375.

Oldenborger GA, Knoll MD, Routh PS, and LaBrecque DJ (2007a) Time-lapse ERT monitoring of an injection/withdrawal experiment in a shallow unconfined aquifer. *Geophysics* 72: F177–F188.

Oldenborger G, Routh P, and Knoll M (2007b) Model reliability for 3D electrical resistivity tomography: Application of the volume of investigation index to a time-lapse monitoring experiment. *Geophysics* 72(4): F167–F175.

Oldenburg DW and Li Y (1994) Inversion of induced polarization data. *Geophysics* 59: 1327–1341.

Oldenburg DW and Li Y (1999) Estimating depth of investigation in DC resistivity and IP surveys. *Geophysics* 64: 403–416.

Olheoft G (1985) Low-frequency electrical properties. *Geophysics* 50: 2492–2503.

Olorunfemi MO and Griffiths DH (1985) A laboratory investigation of the induced polarization of the Triassic Sherwood sandstone of Lancashire and its hydrogeological applications. *Geophysical Prospecting* 33(1): 110–127.

Osiensky J, Nimmer R, and Binley A (2004) Borehole cylindrical noise during hole-surface and hole–hole resistivity measurements. *Journal of Hydrology* 289: 78–94.

Panissod C, Dabas M, Hesse A, Jolivet A, Tabbagh J, and Tabbagh A (1998) Recent developments in shallow-depth electrical and electrostatic prospecting using mobile arrays. *Geophysics* 63(5): 1542–1550.

Parasnis DS (1988) Reciprocity theorems in geoelectric and geoelectromagnetic work. *Geoexploration* 25: 177–198.

Park S and Fitterman D (1990) Sensitivity of the telluric monitoring array in Parkfield, California, to changes of resistivity. *Journal of Geophysical Research* 95: B10.

Park SK and Van GP (1991) Inversion of pole–pole data for 3D resistivity structure beneath arrays of electrodes. *Geophysics* 56: 951–960.

Pelton WH, Sill WR, and Smith BD (1984) Interpretation of complex resistivity and dielectric data, part II. *Geophysical Transactions* 30: 11–45.

Pelton WH, Ward SH, Hallof PG, Sill WR, and Nelson PH (1978) Mineral discrimination and removal of inductive coupling with multifrequency IP. *Geophysics* 43: 588–609.

Portniaguine O and Zhdanov MS (1999) Focusing geophysical inversion images. *Geophysics* 64: 874–887.

Pridmore DF, Hohmann GW, Ward SH, and Sill WR (1981) An investigation of finite-element modeling for electrical and electromagnetic data in three dimensions. *Geophysics* 46: 1009–1024.

Ramirez A, Daily W, Binley A, LaBrecque D, and Roelant D (1996) Detection of leaks in underground storage tanks using electrical resistance methods. *Journal of Environmental and Engineering Geophysics* 1: 189–203.

Ramirez A, Daily W, LaBrecque D, Owen E, and Chesnut D (1993) Monitoring and underground steam injection process using electrical resistance tomography. *Water Resources Research* 29(1): 73–87.

Ren Z and Tang J (2010) 3D direct current resistivity modeling with unstructured mesh by adaptive finite-element method. *Geophysics* 75(1): H7–H17.

Reynolds JM (2011) *An Introduction to Applied and Environmental Geophysics*, 2nd edn. London: Wiley, 696 pp.

Roy A and Apparao A (1971) Depth of investigation in direct current methods. *Geophysics* 36(5): 943–959.

Roy KK and Elliot HM (1980) Resistivity and IP surveys for delineating saline water and fresh water zones. *Geoexploration* 18: 145–162.

Rücker C, Günther T, and Spitzer K (2006) Three-dimensional modelling and inversion of dc resistivity data incorporating topography – I. Modelling. *Geophysical Journal International* 166: 495–505.

Rucker D, Loke M, Levitt M, and Noonan G (2010) Electrical-resistivity characterization of an industrial site using long electrodes. *Geophysics* 75(4): WA95–WA104.

Rucker DF, Noonan GE, and Greenwood WJ (2011) Electrical resistivity in support of geological mapping along the Panama Canal. *Engineering Geology* 117(1–2): 121–133.

Sambuelli L, Comina C, Bava S, and Piatti C (2011) Magnetic, electrical, and GPR waterborne surveys of moraine deposits beneath a lake: A case history from Turin, Italy. *Geophysics* 76(6): B213–B224.

Sasaki Y (1989) Two-dimensional joint inversion of magnetotelluric and dipole–dipole resistivity data. *Geophysics* 54: 254–262.

Sasaki Y (1994) 3D resistivity inversion using the finite-element method. *Geophysics* 59(12): 1839–1848.

Sasaki Y and Matsuo K (1993) Surface-to-tunnel resistivity tomography at the Kamaishi mine. *Batsuri-Tansa* 46: 128–133.

Schenkel CJ (1995) Resistivity imaging using a steel cased well. *Lawrence Livermore National Laboratory Report UCRL-JC-121653*, Livermore, USA.

Schenkel CJ and Morrison HF (1990) Effects of well casing on potential field measurements using downhole current sources. *Geophysical Prospecting* 38: 663–686.

Schima S, LaBrecque DJ, and Lundegard PD (1996) Using resistivity tomography to monitor air sparging. *Ground Water Monitoring and Remediation* 16: 131–138.

Seigel HO (1959) Mathematical formulation and type curves for induced polarization. *Geophysics* 24(3): 547–565.

Seigel HO, Nabighian M, Parasnis DS, and Vozoff K (2007) The early history of the induced polarization method. *Leading Edge* 26: 312–321.

Shima H (1990) Two-dimensional automatic resistivity inversion technique using alpha centers. *Geophysics* 55: 682–694.
Shima H (1992) 2D and 3D resistivity image reconstruction using crosshole data. *Geophysics* 57: 1270–1281.
Simpson D, Van Meirvenne M, Lück E, Bourgeois J, and Rühlmann J (2010) Prospection of two circular Bronze Age ditches with multi-receiver electrical conductivity sensors (North Belgium). *Journal of Archaeological Science* 37(9): 2198–2206.
Slater L and Binley A (2006) Synthetic and field based electrical imaging of a zerovalent iron barrier: Implications for monitoring long-term barrier performance. *Geophysics* 71(5): B129–B137.
Slater L, Binley A, and Brown D (1997a) Electrical imaging of the response of fractures to ground water salinity change. *Ground Water* 35(3): 436–442.
Slater L, Binley A, Daily W, and Johnson R (2000) Cross-hole electrical imaging of a controlled saline tracer injection. *Journal of Applied Geophysics* 44: 85–102.
Slater LD, Brown D, and Binley A (1996) Determination of hydraulically conductive pathways in fractured limestone using cross-borehole electrical resistivity tomography. *European Journal of Environmental and Engineering Geophysics* 1(1): 35–52.
Slater L and Lesmes DP (2002a) Electrical-hydraulic relationships observed for unconsolidated sediments. *Water Resources Research* 38: 1213. http://dx.doi.org/10.1029/2001WR001075.
Slater LD and Lesmes D (2002b) IP interpretation in environmental investigations. *Geophysics* 67(1): 77–88.
Slater LD, Ntarlagiannis D, Day-Lewis FD, et al. (2010) Use of electrical imaging and distributed temperature sensing methods to characterize surface water-groundwater exchange regulating uranium transport at the Hanford 300 Area, Washington. *Water Resources Research* 46: W10533.
Slater L, Zaidman MD, Binley AM, and West LJ (1997b) Electrical imaging of saline tracer migration for the investigation of unsaturated zone transport mechanisms. *Hydrology and Earth System Sciences* 1: 291–302.
Stummer P, Maurer H, and Green AG (2004) Experimental design: Electrical resistivity data sets that provide optimum subsurface information. *Geophysics* 69: 120–139.
Stummer P, Maurer H, Horstmeyer H, and Green AG (2002) Optimization of DC resistivity data acquisition: Real-time experimental design and a new multielectrode system. *IEEE Transactions on Geoscience and Remote Sensing* 40: 2727–2735.
Sumner JS (1976) *Principles of Induced Polarization for Geophysical Exploration*. New York: Elsevier, 277 pp.
Taylor RW and Flemming AH (1988) Characterizing jointed systems by azimuthal resistivity surveys. *Ground Water* 26: 464–474.
Telford WM, Geldart LP, and Sheriff RP (1990) *Applied Geophysics*, 2nd edn. Cambridge, UK: Cambridge University Press.
Tikhonov AN and Arsenin VY (1977) *Solutions of Ill-Posed Problems*. New York: Wiley.
Tsourlos P, Ogilvy RD, Meldrum P, and Williams G (2003) Time-lapse monitoring in single boreholes using electrical resistivity tomography. *Journal of Environmental and Engineering Geophysics* 8: 1–14.
Tsourlos P, Ogilvy RD, Papazachos C, and Meldrum PI (2011) Measurement and inversion schemes for single borehole-to-surface electrical resistivity tomography surveys. *Journal of Geophysics and Engineering* 8: 487–497.
Udphuay S, Günther T, Everett ME, Warden RR, and Briaud J-L (2011) Three-dimensional resistivity tomography in extreme coastal terrain amidst dense cultural signals: Application to cliff stability assessment at the historic D-Day site. *Geophysical Journal International* 185(1): 201–220.
Vacquier V, Holmes R, Kintzinger PR, and Lavergne M (1957) Prospecting for ground water by induced electrical polarization. *Geophysics* 22: 660–687.
Van Schoor M and Binley A (2010) In-mine (tunnel-to-tunnel) electrical resistance tomography in South African platinum mines. *Near Surface Geophysics* 8: 563–574.
Van Voorhis G, Nelson P, and Drake T (1973) Complex resistivity spectra of porphyry copper mineralization. *Geophysics* 38(1): 49–60.
Vernon RW (2008) Alfred Williams, Leo Daft and 'The Electrical Ore-Finding Company Limited'. *British Mining* 86: 4–30.
Versteeg R and Johnson T (2008) Using time-lapse electrical geophysics to monitor subsurface processes. *The Leading Edge* 27: 1488–1497.
Watson KA and Barker RD (1999) Differentiating anisotropy and lateral effects using offset Wenner azimuthal resistivity soundings. *Geophysics* 64: 739–745.
Webster JG (ed.) (1990) *Electrical Impedance Tomography*. New York: Adam Hilger.
Wexler A, Fry B, and Neuman MR (1985) Impedance-computed tomography algorithm and system. *Applied Optics* 24: 3985–3992.
Wilkinson PB, Chambers JE, Kuras O, Meldrum PI, and Gunn DA (2011) Long-term time-lapse geoelectrical monitoring. *First Break* 29: 45–52.
Wilkinson PB, Chambers JE, Meldrum PI, Ogilvy RD, and Caunt S (2006a) Optimization of array configurations and panel combinations for the detection and imaging of abandoned mineshafts using 3D cross-hole electrical resistivity tomography. *Journal of Environmental and Engineering Geophysics* 11: 213–221.
Wilkinson PB, Loke MH, Meldrum PI, et al. (2012) Practical aspects of applied optimized survey design for electrical resistivity tomography. *Geophysical Journal International* 189(1): 428–440.
Wilkinson PB, Meldrum PI, Chambers JE, Kuras O, and Ogilvy RD (2006b) Improved strategies for the automatic selection of optimized sets of electrical resistivity tomography measurement configurations. *Geophysical Journal International* 167: 1119–1126.
Williams KH, Kemna A, Wilkins MJ, et al. (2009) Geophysical monitoring of coupled microbial and geochemical processes during stimulated subsurface bioremediation. *Environmental Science and Technology* 43: 6717–6723.
York T (2001) Status of electrical tomography in industrial applications. *Journal of Electronic Imaging* 10: 608–619. http://dx.doi.org/10.1117/1.1377308.
Yuval and Oldenburg DW (1997) Computation of Cole–Cole parameters from IP data. *Geophysics* 62: 436–448.
Zhang J, Mackie R, and Madden T (1995) 3D resistivity forward modeling and inversion using conjugate gradients. *Geophysics* 60(5): 1313–1325.
Zhou B and Greenhalgh SA (1997) A synthetic study on cross-hole resistivity imaging with different electrode arrays. *Exploration Geophysics* 28: 1–5.
Zhou B and Greenhalgh SA (2001) Finite element three-dimensional direct current resistivity modelling: Accuracy and efficiency considerations. *Geophysical Journal International* 145: 679–688.
Zhou B, Greenhalgh M, and Greenhalgh SA (2009) 2.5-D/3D resistivity modelling in anisotropic media using Gaussian quadrature grids. *Geophysical Journal International* 176: 63–80.
Zohdy A (1989) A new method for the automatic interpretation of Schlumberger and Wenner sounding curves. *Geophysics* 54(2): 245–253.

This page intentionally left blank

11.09 Tools and Techniques: Self-Potential Methods

MD Jackson, Imperial College, London, UK

© 2015 Elsevier B.V. All rights reserved.

11.09.1	**Introduction**	261
11.09.2	**Origin of SP Signals**	262
11.09.2.1	SP Signals of Electrokinetic Origin	262
11.09.2.2	SPs of Electrochemical Origin	265
11.09.2.2.1	Exclusion–diffusion potentials (liquid-junction and membrane potentials)	266
11.09.2.2.2	Redox processes	268
11.09.2.3	SPs of Thermoelectric Origin	269
11.09.2.4	Telluric Currents, Stray Currents, and Sources of Noise in SP Surveys	271
11.09.3	**Data Acquisition**	271
11.09.3.1	Electrode Types	272
11.09.3.2	Laboratory SP Measurements	273
11.09.3.3	Borehole SP Measurements	273
11.09.3.4	Surface SP Measurements	274
11.09.3.4.1	Electrode installation	274
11.09.3.4.2	Base station/reference electrode installation	274
11.09.3.4.3	Survey configurations	275
11.09.4	**Data Interpretation, Modeling, and Inversion**	276
11.09.4.1	Interpretation of Borehole SP Logs	276
11.09.4.2	Interpretation of Field SP Surveys and Monitoring	277
11.09.4.3	SP Forward Modeling Concepts	279
11.09.4.4	Inversion of SP Measurements	280
11.09.5	**Selected Case Studies**	282
11.09.5.1	Mineral Exploration	282
11.09.5.2	Environmental Site Surveying	284
11.09.5.3	Leak Detection from Dams	286
11.09.5.4	Hydrogeology	286
11.09.6	**Future Directions**	288
11.09.7	**Summary**	289
References		290

11.09.1 Introduction

The self-potential (or spontaneous polarization) (SP) method comprises the passive measurement of electric potential at the ground surface and in boreholes. SP measurements are one of the oldest geophysical surveying techniques, dating back to the 1820s when Robert Fox investigated SP anomalies associated with sulfide deposits in a copper mine in Cornwall, the United Kingdom (Fox, 1830). The SP surveying method was developed commercially by the Schlumberger brothers in the 1920s and extended to hydrocarbon reservoir characterization as a borehole logging tool in the 1930s (Schlumberger et al., 1934). Although a relationship between SP and groundwater flow was recognized in the 1930s (e.g., Poldini, 1938), the SP method was primarily applied in mineral exploration and hydrocarbon reservoir characterization until the 1970s, when it was recognized that SP anomalies are also associated with geothermal activity (e.g., Corwin and Hoover, 1979). Since then, applications of the SP method have grown to include environmental and engineering site surveying (e.g., Nyquist and Corry, 2002), hydrology (e.g., Maineult et al., 2008; Revil et al., 2003; Rizzo et al., 2004; Sailhac et al., 2003), monitoring of volcanic and seismic activity (e.g., Aubert and Atangana, 1996; Finizola et al., 2004; Ishido et al., 1989), leakage detection from dams and embankments (e.g., Bogoslovsky and Ogilvy, 1973; Bolève et al., 2009; Ogilvy et al., 1969), location and monitoring of contaminant plumes (e.g., Linde and Revil, 2007; Minsley et al., 2007b; Naudet et al., 2003, 2004), subsurface void and sinkhole detection (e.g., Jardani et al., 2006, 2007), and monitoring of pumping and sparging tests (e.g., Jackson et al., 2012a,b; Maineult et al., 2005; Rizzo et al., 2004). Proposed applications also include monitoring of subsurface flow in hydrocarbon reservoirs (Saunders et al., 2008, 2012) and during CO_2 sequestration (e.g., Moore et al., 2004).

SP methods have a number of advantages over other geophysical techniques: they are often cheaper and quicker to implement, requiring only a pair (or more) of suitable electrodes and a high-impedance voltmeter, and data can be obtained over large regions with dense sampling in both space and time (e.g., Corwin, 1989). Surface SP surveys are rapid and nonintrusive; borehole SP logs are routinely obtained by wireline logging companies. Moreover, SP anomalies are often directly related to the process of interest, such as changes in groundwater flow, chemistry, and/or temperature

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00208-6

(e.g., Corwin and Hoover, 1979; Maineult et al., 2005, 2008). The disadvantages largely lie in interpreting the data, which can be more challenging than other geophysical techniques. Similar to gravity and magnetic methods, SP signals are purely passive, and in many applications, there is no way of adjusting source parameters to help identify signals of interest. Moreover, as this chapter will discuss, SP signals arise from a variety of sources, and distinguishing these can be challenging. Signals that are considered noise in some survey applications (e.g., topography-related streaming potentials in mineral surveying) are the main focus of other monitoring campaigns (e.g., groundwater flow characterization in hydrogeology). SP surveys for mineral exploration, and borehole SP logs, have traditionally been interpreted qualitatively or semiquantitatively; however, a new generation of inversion techniques for SP measurements are now becoming available, driven by improved understanding of the underlying physical processes and increased computing power (e.g., Jardani et al., 2007, 2008; Mendonça, 2008; Minsley et al., 2007a; Woodruff et al., 2010).

11.09.2 Origin of SP Signals

The subsurface charge polarization mechanisms that give rise to SP anomalies have a variety of sources, but a common feature is the presence of water in the pore space of the rock, soil, or sediments. In general, the SP acts to maintain overall electroneutrality when a separation of electric charge occurs in response to natural or induced gradients in thermodynamic potential, such as fluid potential fields (more often expressed as pressure above hydrostatic) and chemical potential (related to oxidation and reduction reactions or variations in ionic concentration) (e.g., Corwin and Hoover, 1979; Marshall and Madden, 1959; Revil, 1999). The local electric current density $\mathbf{j}$ in the ground (A m^{-2}) is given by Ohm's law, plus a current density $\mathbf{j}^i_{sp}$ associated with charge polarization mechanism i

$$\mathbf{j} = -\sigma_s \nabla V + \sum_i \mathbf{j}^i_{sp} \quad [1]$$

where σ_s is the local electric conductivity of the material (saturated rock, soil, sediments, etc., in S m^{-1}) and V is the electric potential (V). Conservation of charge, in the absence of external sources or sinks (such as in resistivity measurements), is expressed as

$$\nabla \cdot \mathbf{j} = 0 \quad [2]$$

Commonly invoked charge polarization mechanisms that give rise to SP anomalies include electrokinetic processes (typically streaming potentials), electrochemical processes (including diffusion, exclusion, and redox potentials), and thermoelectric processes. These are discussed in more detail in the following sections.

11.09.2.1 SP Signals of Electrokinetic Origin

The electrokinetic properties of water-saturated soils, rocks, and unconsolidated sediments are discussed in detail in Section 1.2.6, so only a brief summary is provided here. Electrokinetic phenomena arise as a result of the charge separation that occurs at mineral–water interfaces, to leave an excess of (typically negative) charge on the mineral surfaces and an excess of (typically positive) countercharge in the water adjacent to the mineral surfaces (e.g., Hunter, 1981; Lynch, 1962; Wyllie, 1951 see Section 1.2.4.1). This arrangement of charge at the mineral–water interfaces is known as the electric double layer (or triple layer, depending on the model used to describe the distribution of charge at the interface; see Hunter, 1981; Revil et al., 1999). Separation of charge can also occur at immiscible fluid–fluid interfaces, such as the air–water and oil–water interfaces (e.g., Beattie and Djerdjev, 2004). SPs of electrokinetic origin arise when there is relative motion of the excess charge associated with the materials on each side of the interface.

Streaming potentials are generated when the relative motion of excess charge results from the flow of fluid through a porous medium; in the subsurface, the fluid is typically water containing dissolved salts (**Figure 1**). At the mineral–water interface, there is a so-called shear plane, a small distance from the mineral surface (of order a few nm), at which the water becomes mobile (i.e., can flow relative to the mineral surface); in general, the shear plane is located within the diffuse part of the electric double layer. The electric potential at the location of the shear plane is termed the zeta potential (ζ; see **Figure 1** and Section 1.2.6.1). When water flows through the pore space, some of the countercharge in the diffuse layer is transported by the flow, which gives rise to a streaming current. To maintain overall electroneutrality, a conduction current is established and the electric potential required to maintain the conduction current is the streaming potential. For the flow of a single fluid phase in the absence of other current sources, eqn [1] can be written as

$$\mathbf{j} = -\sigma_s \nabla V_s + \sigma_s C_s \nabla P = -\sigma_s \nabla V_s + Q\mathbf{v} \quad [3]$$

where V_s is the streaming potential (V), C_s is the streaming potential coupling coefficient (V Pa^{-1}), ∇P is the pressure gradient driving the flow (Pa), Q is the excess countercharge transported by the flow (the streaming charge density in C m^{-3}), and $\mathbf{v}$ is the Darcy flow velocity (m s^{-1}). Historically, the streaming current and streaming potential have been described in terms of the coupling coefficient C_s, which was considered to be the key material property of interest (e.g., Revil et al., 2005a; Saunders et al., 2008; Sill, 1983; Vinogradov et al., 2010). The coupling coefficient is a tensor property, but a single scalar value is often defined in a 1-D geometry (such as a capillary tube) when the total current density is zero, which yields from eqn [1]

$$C_s = \left.\frac{\partial V_s}{\partial P}\right|_{j=0} \quad [4]$$

Assuming the double layer is thin relative to the radius of curvature of the pore surfaces (the so-called thin double-layer assumption), the coupling coefficient can be related to the ζ using the well-known Helmholtz–Smoluchowski equation (e.g., Ishido and Mizutani, 1981; see also Section 1.2.6.4)

$$C_s = \frac{\varepsilon \zeta}{\mu \sigma_s F} \quad [5]$$

where ε is the permittivity of the pore water (F m^{-1}), μ is the viscosity of the pore water (Pa s), and F is the (dimensionless)

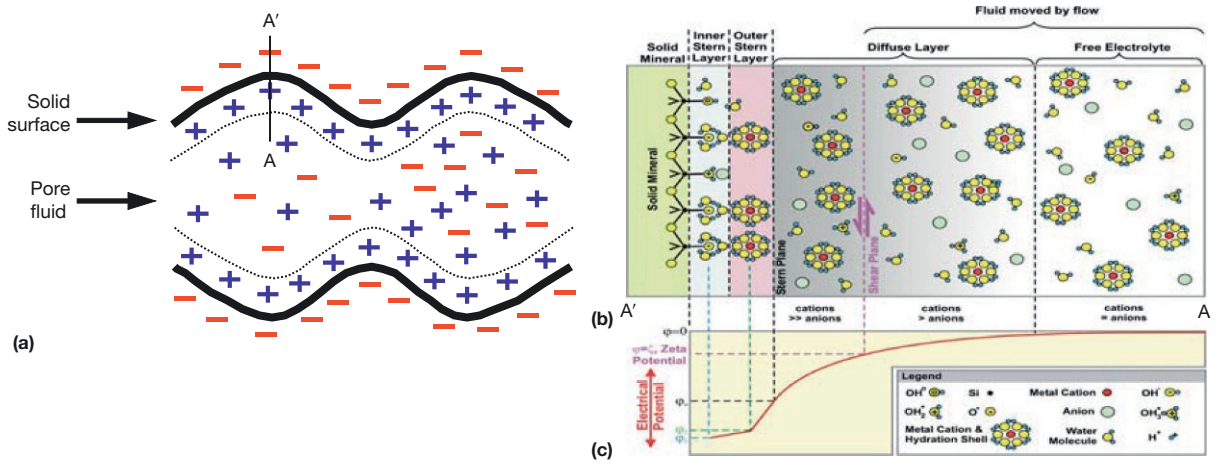


Figure 1 (a) Simple schematic showing the electric double layer within a pore (modified from Jackson MD, Gulamali MY, Leinov E, Saunders JH, and Vinogradov J (2012) Spontaneous potentials in hydrocarbon reservoirs during waterflooding: Application to waterfront monitoring. *Society of Petroleum Engineers Journal* 17: 53–69) and sketches illustrating (b) the distribution of charge for a typical metal oxide such as SiO_2 in contact with water containing dissolved salts and (c) the variation in electric potential, along the profile A–A′ in (a) (from Glover and Jackson, 2010).

formation factor ($=\sigma_w/\sigma_r$ where σ_w is the conductivity of the pore water). The coupling coefficient as defined in eqn [4] is negative if the ζ is negative and vice versa, that is, it has the opposite sign to the streaming charge. Throughout this chapter, the coupling coefficients defined for the various charge polarization mechanisms have the opposite sign to the majority charge carriers.

The thickness of the double layer can be estimated using the Debye length (λ), which varies with salt concentration from *c.* 0.1 nm to a few tens of nm for most natural brines (e.g., **Figure 2(a)**) and is given by (e.g., Section 1.2.4.1)

$$\lambda = \left(\frac{\varepsilon k_B T}{1000 e^2 N_A \sum_j C_j z_j^2} \right)^{1/2} \quad [6]$$

where k_B is the Boltzmann constant ($=1.3806488 \times 10^{-23}$ J K^{-1}), T is temperature (K), e is the charge on an electron ($=1.60217657 \times 10^{-19}$ C), N_A is Avogadro's number ($=6.02214129 \times 10^{23}$ mol^{-1}), C is the molar concentration of ion species j (mol l^{-1}), z is the valency of ion species j, e is the charge on an electron ($=1.60217657 \times 10^{-19}$ C), and the term $\sum C_j z_j^2$ represents the ionic strength of the electrolyte. The advantage of the thin double-layer assumption is that the streaming potential can be expressed in terms of the ζ without having to explicitly solve for the charge distribution in the pore space (see Jackson and Leinov, 2012). The thin double-layer assumption is valid for streaming potential calculations so long as the pore radius is more than approximately 200 times the Debye length; given their relatively large pore size, this is typically the case in earth materials such as sandstones, soils, and unconsolidated sediments, but may not be the case in fine-grained rocks such as mudstones (Jackson and Leinov, 2012). More recently, the streaming potential has been described in terms of the streaming charge density (Q; e.g., Jackson, 2010; Revil and Linde, 2006). In the limit of a thin double layer, Q can be related to ζ via (Jackson, 2010)

$$Q = -\frac{\zeta \varepsilon}{kF} \quad [7]$$

where k is the permeability of the porous material (m^2). The streaming charge density is related to the streaming potential coupling coefficient by

$$C_s = -\frac{Qk}{\sigma_s \mu} \quad [8]$$

Solution of eqns [2] and [3] (or their equivalent integral forms), with the electrokinetic coupling expressed in terms of C_s or Q, lies at the heart of quantitative interpretations of SP anomalies resulting from streaming potentials (see **Section 11.09.4.4**).

Equation [3] shows there is a direct relationship between streaming potential and fluid flow, which makes SP an attractive method for detecting and monitoring flow in the subsurface. SP anomalies caused by streaming potentials are observed when natural groundwater flow occurs in response to rainfall and topography (e.g., Doussan et al., 2002; Perrier and Morat, 2000; Trique et al., 2002), seawater intrusion in coastal areas (e.g., Revil et al., 2005b), and convection in geothermal areas and volcanoes (e.g., Aubert and Atangana, 1996; Finizola et al., 2004; Ishido et al., 1989; Michel and Zlotnicki, 1998); moreover, the SP anomalies may be focused on specific flow paths, allowing identification of subsurface features such as sinkholes, faults, and karsts (e.g., Freund et al., 2007; Jardani et al., 2006, 2007). SP anomalies are also induced during groundwater abstraction, by leakage from dams and embankments, and by water injection into geothermal reservoirs and hydrocarbon reservoirs (e.g., Bogoslovsky and Ogilvy, 1973; Darnet et al., 2004; Maineult et al., 2008; Marquis et al., 2002; Rizzo et al., 2004; Saunders et al., 2006, 2008; Titov et al., 2000). However, the magnitude of the streaming potential that results from a given pressure gradient is strongly dependent on a number of factors, including the salinity of the pore water, the pH, the temperature, and the surface properties of the minerals comprising the porous geomaterial (see Section

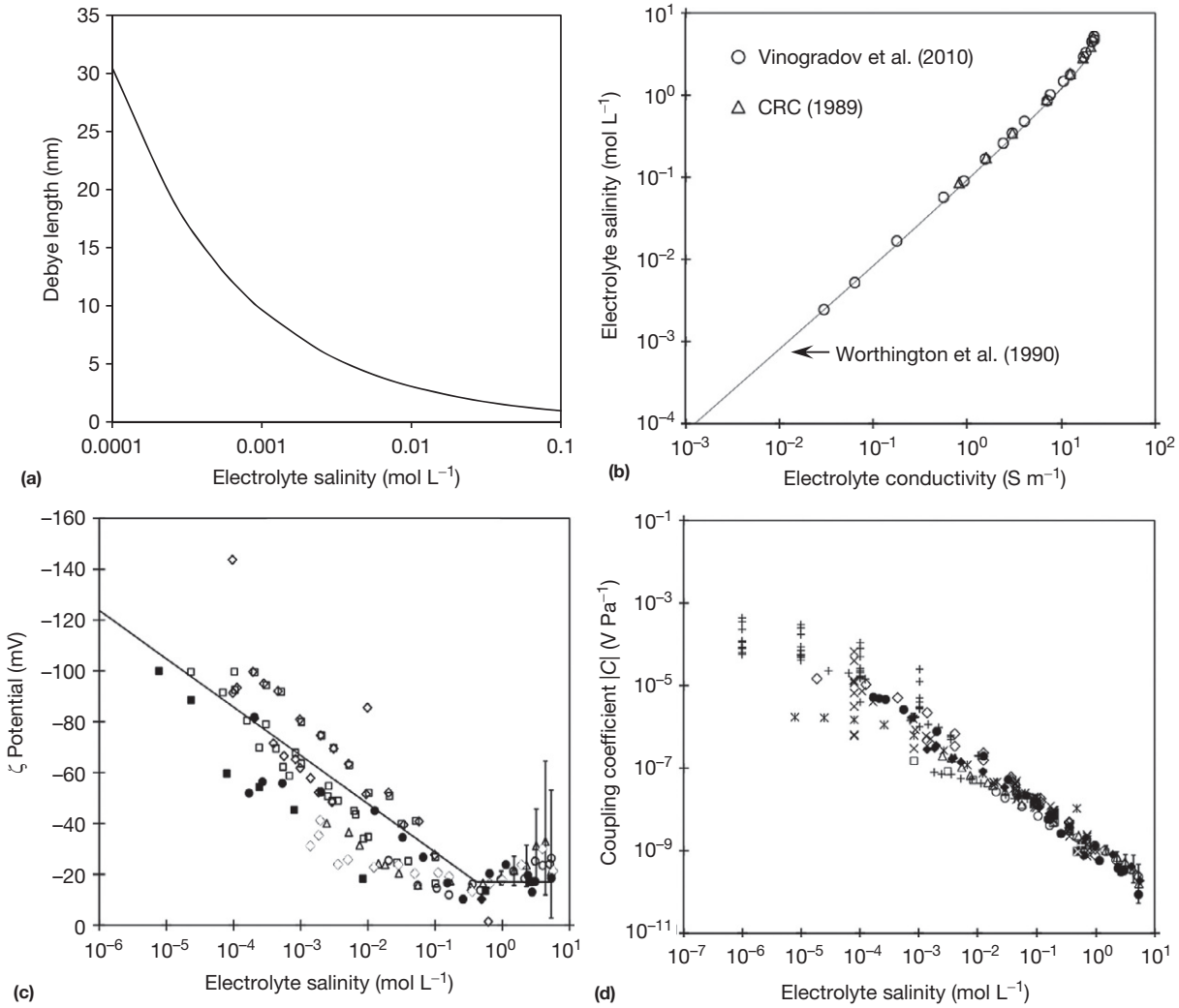


Figure 2 Salinity dependence of (a) Debye length, calculated using eqn [6] and assuming a simple symmetrical monovalent electrolyte and constant permittivity, (b) electric conductivity of NaCl brine, and (c) ζ potential for sandstone and quartz saturated with NaCl brine; the regression is given by eqn [9a] and [9b], and (d) streaming potential coupling coefficient. Plots (b)–(d) modified from Vinogradov J, Jaafar MZ, and Jackson MD (2010) Measurement of streaming potential coupling coefficient in sandstones saturated with natural and artificial brines at high salinity. *Journal of Geophysical Research* 115: B12204. http://dx.doi.org/10.1029/2010JB007593.

1.2.6.7 for a complete list); these factors dictate the values of σ_r, C_s, and/or Q_s in eqn [3].

Streaming potentials are strongly dependent on salinity for two reasons. First, the electric conductivity of the pore water is salinity-dependent, increasing with increasing salinity (e.g., **Figure 2(b)**); as the conductivity increases, the streaming potential required to drive the conduction current that ensures overall electroneutrality decreases. Second, the thickness of the diffuse part of the electric double layer decreases with increasing salinity (**Figure 2(a)**); however, the location of the shear plane remains approximately constant, so, as the double layer thickness decreases, the ζ (and the streaming charge density) also decreases (**Figure 2(c)**). The salinity dependence of electric conductivity and ζ means that larger streaming potentials for a given pressure gradient are observed at low salinity and vice versa (**Figure 2(d)**). Vinogradov et al. (2010) proposed an empirical relationship between salinity and ζ in quartz and quartz-dominated sandstones saturated with NaCl brine at 25 °C and pH 7 given by

$$\zeta = a + b \log C_{NaCl} \quad C_{NaCl} < 0.4\,M \qquad [9a]$$

$$\zeta = c \quad C_{NaCl} \geq 0.4\,M \qquad [9b]$$

where the experimentally determined constants are $a = -9.67$ mV, $b = 19.02$ mV, and $c = -17$ mV; the ζ is expressed in mV; and the salinity C_{NaCl} is expressed in mol l^{-1}. A relationship of the form of eqn [9a] between ζ and salinity was originally proposed based on measured ζ data in the low salinity range (<0.1 M) (Pride and Morgan, 1991) and can also be derived from a model of the electric double layer, which captures the partitioning of countercharge between the Stern and diffuse layers and invokes the Boltzmann equation to describe the charge density in the diffuse layer (Revil et al., 1999). More sophisticated relationships that account for the

temperature, pH, ionic species dissolved in the brine, and the surface properties of the minerals have also been developed from models of the double layer (e.g., Section 1.2.6.6; see also Glover et al., 2012), although the limited experimental data available mean that their predictions have not yet been thoroughly validated. In particular, the behavior of the ζ at elevated temperature is still poorly understood (e.g., Reppert and Morgan, 2003; Saunders et al., 2012).

The streaming potential also depends on the fluid phases present in the pore space and has been shown to be dependent on water saturation. If water is the only charge-carrying fluid phase (i.e., $Q = Q_w$ where Q_w is the streaming charge density in the water), which is typically the case if the second fluid phase is a gas (such as air) or a nonaqueous phase liquid such as oil, then experimental data suggest that the dependence of the streaming charge density on water saturation during drainage is described empirically by a relationship of the form

$$\frac{Q_w(S_w)}{Q_w(S_w=1)} = Q_{rw}(S_w) = d\left(\frac{S_w - S_{wirr}}{1 - S_{wirr}}\right)^e + f \qquad [10]$$

where S_{wirr} is the minimum water saturation at the end of drainage; d, e, and f are constants determined from experimental data; and $d+f=1$ (**Figure 3**). The multiphase streaming potential coupling coefficient is related to the multiphase streaming charge density by (Jackson, 2010)

$$\frac{C_s(S_w)}{C_s(S_w=1)} = C_{rs}(S_w) = k_{rw}(S_w)Q_{rw}(S_w)I(S_w) \qquad [11]$$

where k_{rw} is the water relative permeability and $I(S_w)$ is the resistivity index ($\sigma_s(S_w=1)/\sigma_s(S_w)$; see, e.g., Section 1.2.2.3).

Values of C_s for earth materials typically have magnitudes in the range $|10^{-5}|$–$|10^{-10}|$ V Pa^{-1}, with corresponding values of ζ in the range $|150|$–$|10|$ mV and Q in the range $|1|$–$|10^5|$ C m^{-3} (see, e.g., **Figure 2** and also **Figure 22**, Section 1.2.6). The coupling coefficient and ζ are typically negative for silicate minerals but may be positive for carbonates, depending upon the salts dissolved in the groundwater (see, e.g., Strand et al., 2006; Zhang and Austad, 2006). The magnitude of the corresponding streaming potential may vary from μV to 100s mV, depending upon the wide variety of factors described earlier. Thus, even when it is clear that streaming potentials dominate a measured SP anomaly, quantitative interpretation of the data can be challenging (see **Section 11.09.4**).

Note that the coupling between flow and electric potentials is symmetrical in linear thermodynamics. Applying an electric potential to a charged, water-saturated porous medium such as a rock results in fluid flow (a process termed electro-osmosis; Hunter, 1981); likewise, applying an electric potential to a suspension of charged particles results in motion of the particles (electrophoresis; Hunter, 1981). The former process has been applied in a technique termed 'electrokinetic remediation' for the in situ removal of ionic inorganic compounds, such as heavy metals. Electrokinetic remediation imposes electric currents of the order mA cm^{-2}, and the dominant removal mechanisms are electro-osmosis and ionic migration, in which the electric field induces movement of ions (e.g., Virkutyte et al., 2002). The latter process is applied in commercial zetameters, which are used to measure the ζ of charged particles and colloids in suspension; see Kosmulski and Dahlsten (2006) for typical experimental data on geomaterials. Electrokinetics also lie at the heart of the so-called seismoelectric effect, in which the passage of a seismic wave causes relative motion of the pore water, inducing an alternating (AC) streaming current and associated SP response (e.g., Thompson and Gist, 1993; see also Section 1.2.6.3).

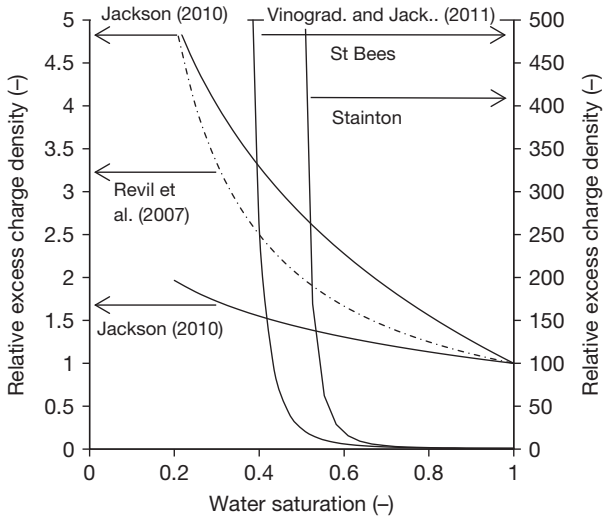


Figure 3 Saturation dependence of the relative streaming charge density Q_{rw} during drainage fitted to (i) the model of Revil and coworkers that assumes $Q_{rw} = 1/S_w$ (e.g., Revil et al., 2007; dot-dash line, modeled with $d=1$, $e=-1$, $f=0$, and $S_{wirr}=0$ in eqn [11], curve truncated at a minimum water saturation of 0.2 and plotted against the left-hand axis), (ii) the capillary tubes model of Jackson (2010) with varying capillary size distributions (solid lines modeled with (lower) $d=600$, $e=-0.001$, $f=-599$ and (upper) $d=2500$, $e=-0.001$, $f=-2499$, $S_{wirr}=0$, curves truncated at a minimum water saturation of 0.2 and plotted against the left-hand axis), and (iii) the experimental data of Vinogradov and Jackson (2011) (dashed lines modeled with (St Bees) $d=0.25$, $e=-3.7$, $f=0.75$, $S_{wirr}=0.29$, and (Stainton) $d=0.22$, $e=-3.5$, $f=0.78$, $S_{wirr}=0.44$, both curves plotted against the right-hand axis). St Bees and Stainton denote the sandstone samples investigated by Vinogradov and Jackson (2011). Note the experimental data yields much larger Q_{rw} than the simple models of Revil et al. (2007) and Jackson (2010).

11.09.2.2 SPs of Electrochemical Origin

SPs of electrochemical origin arise when charge transport occurs in response to variations in chemical potential. In the subsurface, variations in chemical potential may result from variations in the concentration of salt species in the water occupying the pore space or variations in redox potential associated with, for example, variations in oxygen fugacity or pH. SP signals arising from redox processes are invoked to explain the so-called mineral potentials associated with ore deposits (e.g., Bigalke and Grabner, 1997; Mendonça, 2008; Sato and Mooney, 1960; Stoll et al., 1995; Telford et al., 1990) and also the SP signals associated with contaminant plumes (e.g., Arora et al., 2007; Linde and Revil, 2007; Minsley et al., 2007b; Naudet et al., 2003, 2004). SP signals arising from concentration gradients, which are termed here exclusion–diffusion

potentials, are responsible for the variations in SP associated with permeable bed boundaries in borehole logs (e.g., Telford et al., 1990; Woodruff et al., 2010).

11.09.2.2.1 Exclusion–diffusion potentials (liquid-junction and membrane potentials)

Variations in water composition result in concentration gradients down which the salt species migrate. However, the salt ions have differing mobility so do not migrate at the same rate. For example, the mobility of sodium ions at 25 °C is around 70% that of chloride ions (e.g., Braun and Weingartner, 1985). The resulting charge separation is countered by an exclusion–diffusion (or liquid-junction) potential (V_{ed}), to maintain electroneutrality (e.g., Revil, 1999). For the diffusion of ions in response to a concentration gradient ∇C within a porous medium saturated with a dilute, monovalent, symmetrical electrolyte, eqn [1] can be written as (e.g., Revil and Linde, 2006)

$$\mathbf{j} = -\sigma_s \nabla V_{ed} + \sigma_s C_{ed} \nabla \ln C \qquad [12]$$

where C_{ed} is the exclusion–diffusion coupling coefficient (in $V\,e^{-1}$, where e is the natural logarithm). For a porous medium saturated with a dilute, monovalent, symmetrical electrolyte, the coupling coefficient can be written as (Revil, 1999)

$$C_{ed} = -\frac{k_B T}{e}(T_+ - T_-) \qquad (13)$$

where $T_\pm$ are the macroscopic (dimensionless) Hittorf transport numbers of the salt ions in the porous media and

$$T_+ + T_- = 1 \qquad [14]$$

If the solid surfaces are not electrically charged, the SP is solely due to the diffusion potential, which maintains electroneutrality by retarding the migration of high-mobility ionic species and enhancing the migration of low mobility species; the diffusion potential is the electric potential required to maintain electroneutrality when the diffusion of ions results in charge transport (e.g., Revil, 1999). In this case, the macroscopic Hittorf numbers $T_\pm$ are equivalent to the microscopic values $t_\pm$, given by

$$t_\pm = \frac{\beta_\pm}{\beta_+ + \beta_-} \qquad [15a]$$

$$t_+ + t_- = 1 \qquad [15b]$$

where $\beta_\pm$ represents the mobility of the salt ions within the electrolyte ($m^2\,V^{-1}\,s^{-1}$) and the exclusion–diffusion potential reduces to the diffusion potential V_d, which, for a concentration difference $C_1 - C_2$ (with $C_1 > C_2$ and C in $mol\,l^{-1}$) in a dilute, monovalent, symmetrical electrolyte, with the reference voltage level taken at concentration C_2, can be expressed as (e.g., Revil et al., 2005a,b)

$$V_d = -\frac{k_B T}{e}(t_+ - t_-)\ln\left(\frac{C_1}{C_2}\right) = -\frac{k_B T}{e}\left(\frac{\beta_+ - \beta_-}{\beta_+ + \beta_-}\right)\ln\left(\frac{C_1}{C_2}\right) \qquad [16]$$

However, the presence of an electric double layer at the mineral–water interface (Figure 1) means that, in water-saturated geomaterials such as rocks, soils, or sediments, some of the (typically negative) co-ions in the brine are excluded from the pore space, so a net excess of (typically positive) countercharge migrates down the concentration gradient. This gives rise to an exclusion (or membrane) potential (V_e), which, for a concentration difference $C_1 - C_2$ (with $C_1 > C_2$) across a negatively charged porous medium, from which all co-ions have been excluded from the pore space (i.e., a perfect ion-selective membrane), and with the reference voltage at concentration C_2, is given by (e.g., Revil et al., 2005a)

$$V_e = -\frac{k_B T}{e}\ln\left(\frac{C_1}{C_2}\right) \qquad [17]$$

The sign of eqn [17] is reversed if the porous medium is positively charged, which may be the case in carbonates (e.g., Jackson and Vinogradov, 2012). If the dominant salt species is NaCl, then $t_{Cl}/t_{Na} \approx 1.49$; hence, eqns [16] and [17] become

$$V_d = 0.021\,T\ln\left(\frac{C_1}{C_2}\right) \qquad [18]$$

$$V_e = -0.086\,T\ln\left(\frac{C_1}{C_2}\right) \qquad [19]$$

where V_d and V_e are in mV (Figure 4). Note that diffusion potentials in NaCl electrolyte and exclusion potentials across negatively charged porous media must have opposing polarity, because the dominant charge carriers are anions and cations, respectively.

Equations [16] and [17] can also be found written in a variety of forms. Often, the term $k_B T/e$ is replaced by the equivalent term RT/F, where R is the molar gas constant ($=8.3144621\,J\,K^{-1}\,mol^{-1}$) and F is the Faraday constant ($=9.6483399 \times 10^4\,C\,mol^{-1}$) (e.g., Jordan and Campbell, 1986; Telford et al., 1990); the electrolyte concentration C may also be replaced by the activity a (e.g., Jordan and Campbell, 1986) or electric conductivity σ_f (e.g., Revil, 1999). The use of activity in eqns [16] and [17] accounts for the nonideality of the electrolyte; the activity is related to concentration by $a = \gamma C$, where γ is the activity coefficient. Tabulated values of activity coefficient are available for many common salt species and typically approach unity for monovalent salt species such as Na^+ and Cl^- at low concentration and/or temperature (Figure 5). Electric conductivity or resistivity may also be used as a proxy for salinity, assuming there is a close-to-linear relationship between conductivity and activity or salinity, for a given temperature and electrolyte composition (e.g., Revil, 1999).

The total SP arising from concentration gradients in a charged porous medium is the sum of the contributing diffusion and exclusion potentials. However, this does not mean that the total SP is given by the sum of eqns [16] and [17] as is claimed in some textbooks (e.g., Telford et al., 1990, p. 284, eqn [5.4]). This assumption arises from an erroneous generalization of the exclusion–diffusion polarization mechanisms that give rise to the SP response in borehole logs (see Section 2.6.3.1). Equations [16] and [17] describe the end-members when either the diffusion potential (i.e., there is no exclusion of charge from the pore space, so the exclusion potential is zero) or the exclusion potential (i.e., all co-charge has been excluded from the pore space, so only the countercharge is able to migrate down the concentration gradient) dominates. In

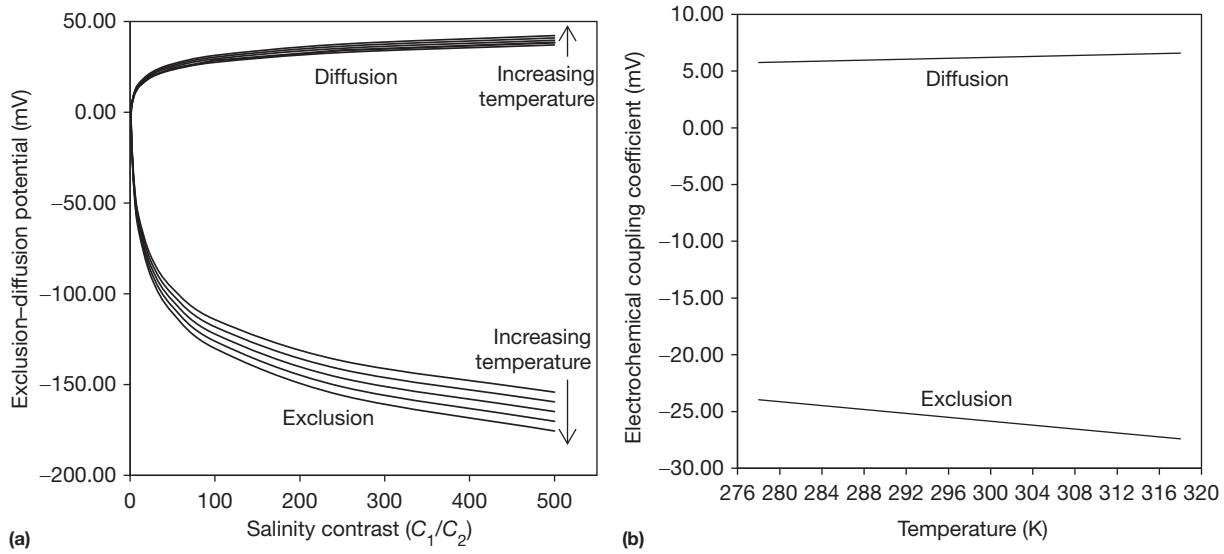


Figure 4 Limiting values of (a) exclusion and diffusion potentials for a dilute, symmetrical, monovalent electrolyte as a function of salinity contrast C_1/C_2, calculated using eqns [16] and [17] over the temperature range 15–55 °C (288–328 K), and (b) the exclusion–diffusion coupling coefficient C_{ed}, for the exclusion and diffusion dominated cases, as a function of temperature, calculated using eqn [12]. In all cases, $t_{Na} = 0.39$ and the relative mobility is assumed constant and independent of temperature and salinity, which is reasonable over the range shown.

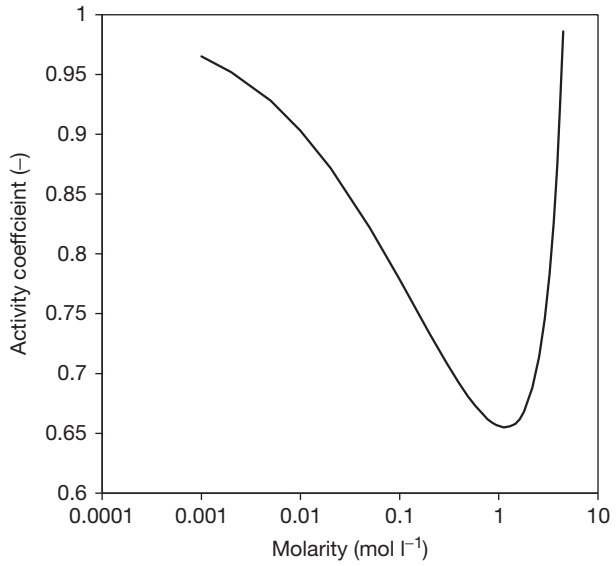


Figure 5 Mean activity coefficients for NaCl electrolyte as a function of molarity. Reproduced from Hamer WJ and Wu YC (1972) *Journal of Physical and Chemical Reference Data* 1: 1047–1099.

general, V_{ed} across a charged porous medium is given by (Revil et al., 2005a)

$$V_{ed} = -\frac{k_B T}{e}(T_+ - T_-)\ln\left(\frac{C_1}{C_2}\right) \quad [20]$$

If diffusion potentials dominate, then $T_\pm = t_\pm$; if exclusion potentials dominate, then $T_\pm = 1$, with $T_+ = 1$ in a negatively charged membrane and $T_- = 1$ in a positively charged membrane. Between these limiting cases, the values of $T_\pm$ are dictated by the relative contribution of the diffusion and exclusion potentials to the overall SP. This depends upon the mobility contrast between ionic species, the thickness of the electric double layer relative to the radius of the water-occupied pores, and the surface charge of the porous medium (e.g., Ortiz et al., 1973; Revil, 1999; Revil et al., 2005a; Westermann-Clark and Christoforou, 1986). If the double layer is thin compared to the pore radius, the ionic mobility contrast is large, or the surface charge on the porous media is small, then the diffusion potential dominates, and vice versa. Given their relatively large pore size relative to the thickness of the double layer, charge exclusion is negligible in earth materials such as sandstones, soils, and unconsolidated sediments, so diffusion potentials typically dominate; however, exclusion potentials may dominate in rocks such as mudstones (the 'shale potential' in SP logging applications), because these are fine-grained so have a small pore size and a large surface electric charge. Exclusion–diffusion potentials may vary from μV to 100s mV (e.g., **Figure 4**), but must lie between the end-members of pure diffusion and pure exclusion.

The most widespread application of SP arising from diffusion–exclusion processes is the response at permeable bed boundaries observed in borehole logging tools. This is primarily caused by diffusion potentials arising from variations in salt concentration between drilling fluids and groundwater, and exclusion potentials across fine-grained beds such as mudstones. The measured SP is assumed to be the difference between these diffusion and exclusion potentials; if the salt in the drilling fluid and groundwater is predominantly NaCl, and the mudstones exclude all the Cl^- anions, the SP is then given by the difference between eqns [16] and [17] (see **Section 11.09.4.1**). Exclusion–diffusion potentials also give rise to SP anomalies when saline water intrudes freshwater (e.g., Martinez-Pagan et al., 2010), saline water is injected as a tracer (e.g., Ikard et al., 2012), or relatively fresh water is injected into oil reservoirs containing more saline brine

(e.g., Gulamali et al., 2011). SP anomalies related to exclusion–diffusion potentials are also created by biological activity; for example, plant roots are ion-selective membranes, and SP signals of order hundreds of mV can be related to changes in vegetation (e.g., Nyquist and Corry, 2002; Telford et al., 1990) or microbial activity (e.g., Atekwana et al., 2006).

11.09.2.2.2 Redox processes

SP anomalies of electrochemical origin arise when gradients in redox potential result in charge separation. In contrast to the exclusion and diffusion potentials described in the preceding text, the polarization mechanisms associated with gradients in redox potential are relatively poorly understood. However, given that oxidation and reduction reactions involve the donation and reception of electrons, respectively, it is clear that electrons must act as charge carriers in some part of the process. SP anomalies associated with redox processes are often large, reaching several 100s or 1000s of mV over ore bodies (e.g., Sato and Mooney, 1960; Telford et al., 1990 and references therein), buried metallic objects such as borehole casings and pipes (e.g., Bigalke and Grabner, 1997; Castermant et al., 2008; Hubbard et al., 2003; Rittgers et al., 2013), and contaminant plumes (e.g., Linde and Revil, 2007; Minsley et al., 2007b; Naudet et al., 2003, 2004; Rittgers et al., 2013). In the former two cases, the route for electron transport is clear: the electrically conductive ore body or metallic object. However, in the latter case, the transport path for electrons is less obvious; such electronic conduction paths may result from (i) the presence of metallic precipitates, which are the end products of the biodegradation of oil, and/or (ii) interconnected, cellular nanowire networks (e.g., Fachin et al., 2012; Gorby et al., 2006; Kato et al., 2012; Ntarlagiannis et al., 2007; Revil et al., 2010). Gradients in redox potential may be associated with depth (differing, e.g., above and below the water table), contaminant concentration in groundwater, aerobic and anaerobic microbial activity, nutrient availability, and variations in pH and oxygen fugacity relating to any of these factors (Atekwana et al., 2006; Doherty et al., 2010; Minsley et al., 2007b; Naudet et al., 2003, 2004; Nyquist and Corry, 2002; Revil et al., 2010; Sato and Mooney, 1960; Stoll et al., 1995; Timm and Möller, 2001).

Sato and Mooney (1960) suggested that the SP anomalies associated with ore bodies (often termed mineral potentials) arise in response to gradients in redox potential in the surrounding water-saturated rock, from relatively oxidizing conditions near the top of the ore body (above the water table) to relatively reducing conditions near the base of the ore body (below the water table). Migration of ions through the groundwater, and electrons through the ore body, constitutes an electric current circuit, in an arrangement often termed a 'geobattery' (e.g., **Figure 6(a)**). Oxidized species in the groundwater approach the top of the ore body and receive electrons and vice versa at the base; consequently, the top of the ore body acts as the cathode of the 'geobattery' (i.e., as an electron source) and the base of the ore body acts as the anode. Buried metallic objects that span a gradient in redox potential behave in a similar way (e.g., Castermant et al., 2008; Rittgers et al., 2013). Note that SP surveys do not record the redox potential directly. Rather, it is the gradient in redox potential within the groundwater, and associated transport of electrons through the ore body or some other electrically conductive path, that gives rise to charge transport in the groundwater; SP surveys record the associated SP anomaly (Bigalke and Grabner, 1997; Linde and Revil, 2007; Stoll et al., 1995). The requirements for an SP anomaly associated with a subsurface redox potential gradient are (i) a good electronic conductor that spans the gradient in redox potential and (ii) electron exchange between the electronic conductor and ions in the neighboring groundwater (Revil et al., 2010; Rittgers et al., 2013). Note that these general criteria encompass the ore body-specific criteria suggested by Sato and Mooney (1960).

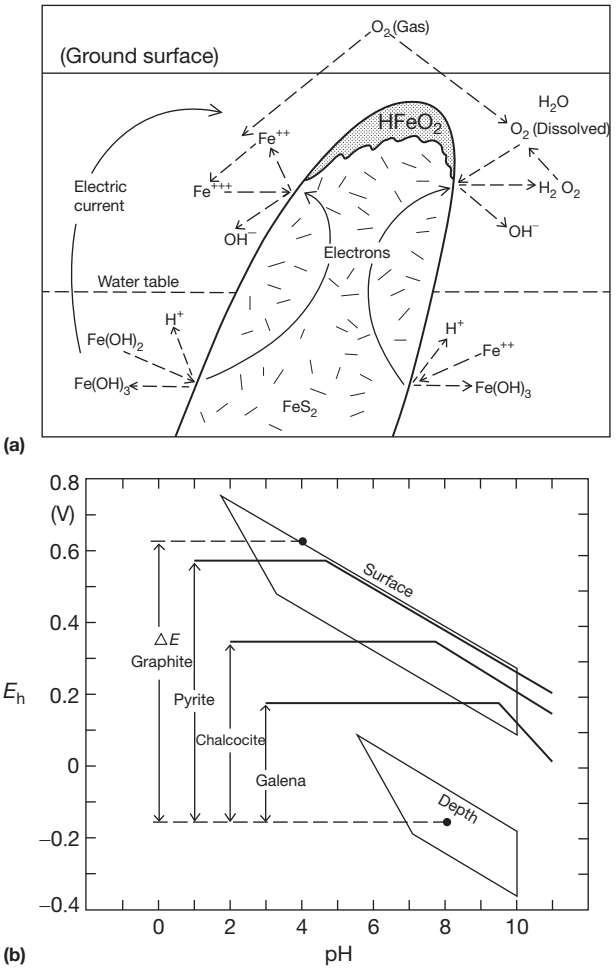


Figure 6 (a) Mineral potential model of Sato and Mooney (1960) and (b) estimated maximum mineral potential using this model (reproduced from Sato M and Mooney HM (1960) The electrochemical mechanism of sulfide selfpotentials. *Geophysics* 25: 226–249).

In the Sato and Mooney model, the mineral potential associated with an ore body can be estimated by treating the oxidation and reduction reactions as two 'half-cell reactions,' each with its own redox potential. Conventionally, this potential is defined with reference to the oxidation potential of hydrogen (i.e., the reaction $H_2 \rightarrow 2H^+ + 2e^-$) in which case the half-cell potential is denoted E_h. The value of E_h for a given oxidation–reduction reaction is then given by (Sato and Mooney, 1960; Timm and Möller, 2001)

$$E_h = E_h^0 - \frac{k_B T n_H}{e n_e} \mathrm{pH} + \frac{k_B T n_p}{e n_e n_r} \log\left(\frac{a_p}{a_r}\right) \quad [21]$$

where E_h^0 is the standard oxidation potential (V) for the reaction (corresponding to unit activities of all chemical species); a_r and a_p are the activities of the reactant and product, respectively; and n_e, n_H, n_r, and n_p denote the mole numbers of participating electrons, hydrogen ions, reactants, and products, respectively. The reaction controlling E_h in the oxidizing region, close to the surface, involves the reduction of O_2 (see also Rittgers et al., 2013). At depth, in the reducing region, Sato and Mooney (1960) estimated upper and lower bounds for E_h based on the stability of siderite and hematite, given that these are often observed to be present in unweathered rocks and ore deposits. The maximum E_h between oxidizing and reducing zones is given by the sum of the E_h values for the respective reactions; assuming pH 4 and strongly oxidizing conditions above the water table, assuming pH 8 below the water table, and taking the average E_h for siderite and hematite, Sato and Mooney (1960) estimated a maximum E_h from the top to the base of an ore body of 0.78 V (e.g., **Figure 6(b)**). They argued that this maximum value applies to inert ore bodies such as graphite, which oxidize at negligible rates in an aqueous environment. If there is appreciable oxidation of the ore body, the E_h associated with this reaction acts to oppose the oxidation E_h, yielding a lower total E_h from the top to the base of the ore body. Sato and Mooney concluded that the mineral potential associated with an ore body, assuming a reference electrode located at infinity, is approximately half the E_h value shown in **Figure 6(b)**; consequently, mineral potential anomalies should not exceed *c.* 400 mV. However, SP anomalies associated with ore bodies of >1 V have been recorded, which exceeds the maximum predicted by the Sato and Mooney model.

Stoll et al. (1995) developed a model for the SP associated with inert ore bodies (the inert electrode model (IEM) of Mendonça, 2008), which not only accounts for the gradient in E_h with depth predicted by Sato and Mooney (1960) but also includes electrode kinetics at the mineral surface. The specific reactions giving rise to the E_h gradient were not considered. Stoll et al. (1995) noted that, at equilibrium for an inert ore, there is no net flow of current across the ore–groundwater interface, so no potential drop can be measured in the groundwater; instead, the problem must be described in terms of the current exchange density across the interface. At the interface, the conduction mechanism changes from electronic to electrolytic; the current exchange density therefore represents the source density for the ionic current within the groundwater. Equation [1] becomes

$$\mathbf{j} = -\sigma_s \nabla V - j_e \frac{nF}{RT}(E_h - E_m + V) \quad [22]$$

where j_e is the exchange current density, which is zero everywhere except at surface of the ore body; E_m is the electric potential of the ore body, which is assumed constant owing to the high electric conductivity of the body; E_h is the (varying) redox potential in the groundwater; and n is the number of electrons exchanged per reaction at the interface. The term $RT/j_e nF$ is called the polarization resistance (units of $\Omega\,m^2$). Stoll et al. (1995) solved eqn [22] numerically, subject to the constraint of eqn [2], and, using a trial and error method and assuming a constant but adjustable value of the polarization resistance, were able to match the SP anomaly measured at the surface and in a nearby borehole associated with massive graphitic mineralization in northeast Bavaria, Germany (**Figure 7**). The model of Stoll et al. (1995) was applied by Mendonça (2008) in both forward and inverse modeling of SP associated with inert ore bodies; inverse modeling allows the magnitude and location of the exchange current sources at the surface of an ore body to be identified from measured SP data, assuming a known subsurface conductivity structure. The approach, discussed in more detail in **Section 11.09.4.4**, yielded a good match to published data from two gold deposits in the Yanacocha district of Peru (Goldie, 2002).

In a different approach, initially developed to interpret SP measurements associated with contaminant plumes and later extended to the general case, Naudet et al. (2003) suggested that, when charge transport occurs in response to gradients in redox potential, eqn [1] can be expressed as

$$\mathbf{j} = -\sigma_s \nabla V - \sigma_c C_h \nabla E_h \quad [23]$$

where C_h is a dimensionless sensitivity coefficient, analogous to those in eqns [3] and [12], and σ_c is the electric conductivity of the region or material spanning the redox potential gradient (Castermant et al., 2008; Linde and Revil, 2007; Rittgers et al., 2013). At a contaminated landfill site near Marseille, Naudet et al. (2003) and Arora et al. (2007) estimated a constant value of $C_h = 0.5$, by removing the streaming potential component from SP survey data and correlating the residual SP signal with direct measurements of redox potential. Naudet et al. (2004) commented on the close match between this value of C_h and the model of Sato and Mooney (1960) in which the SP anomaly is expected to be half of the E_h value. However, this is the only interpreted value of C_h to date; later studies have taken $C_h = 1$ (Castermant et al., 2008; Rittgers et al., 2013). Minsley et al. (2007b) inverted SP data measured at a site contaminated with dense nonaqueous phase liquids to identify the spatial distribution of current sources (i.e., the second term on the right-hand side of eqn [1]); they then assumed a simple linear relationship between the current source density and contaminant concentration. However, the match between observed and modeled contaminant concentrations was of variable quality. This may be because streaming potentials also contributed to the measured SP data; it is also likely that the redox potential does not correlate directly with contaminant concentration but also contaminant type, oxygen fugacity, and microbial activity.

11.09.2.3 SPs of Thermoelectric Origin

Thermoelectric potentials arise when charge transport occurs in response to variations in temperature; the temperature gradient creates a gradient in chemical potential, down which ionic species migrate (e.g., Leinov et al., 2010; Revil, 1999). However, as discussed in **Section 11.09.2.2.1**, the ions have differing mobility so do not migrate at the same rate. Moreover, the presence of an electric double layer at the mineral–water interface means that, in water-saturated geomaterials such as rocks, soils, or sediments, some of the (typically negative) co-ions in the brine are excluded from the pore space, so a net excess of

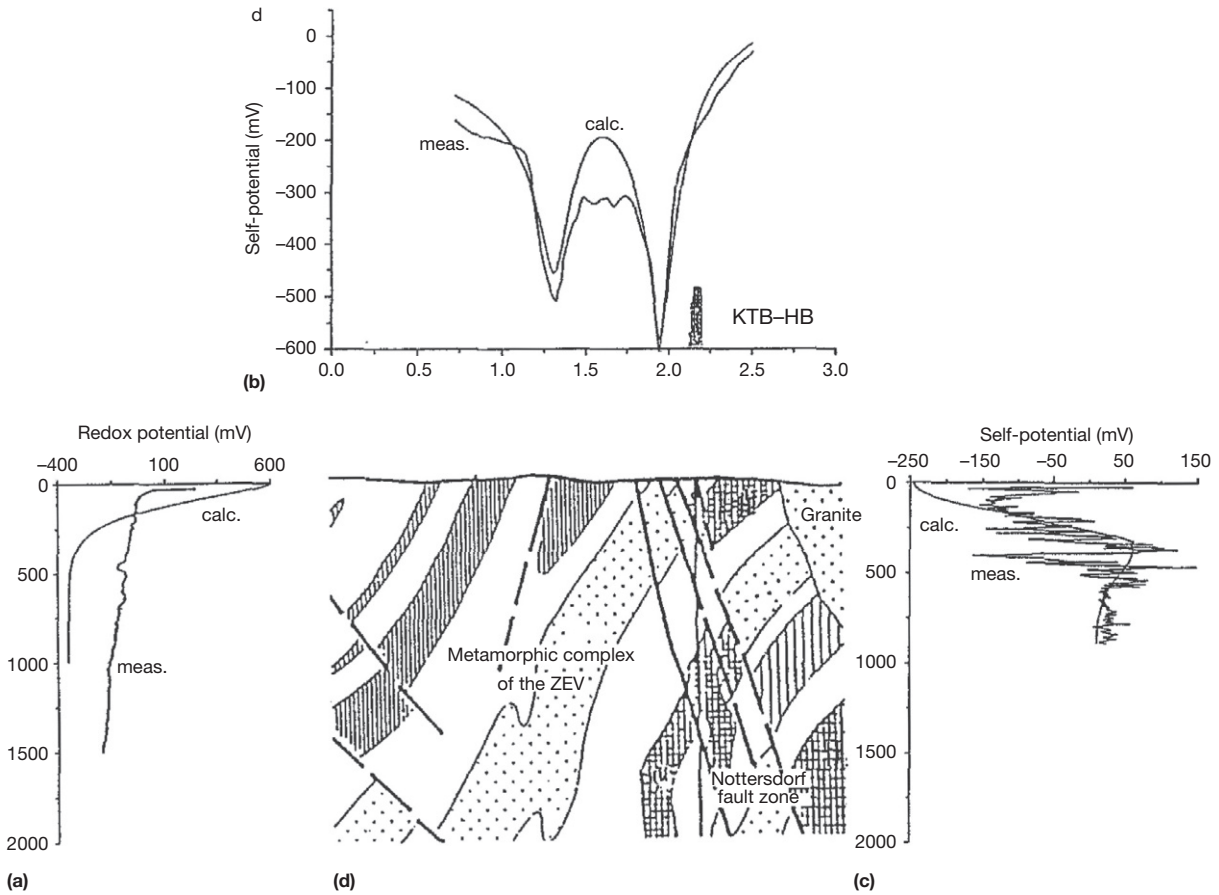


Figure 7 Match between (a) redox potential measured (denoted 'meas.') downhole, (b) SP anomaly measured at the surface, and (c) SP anomaly measured downhole, and predictions (denoted 'calc.') using the model of Stoll et al. (1995). The SP anomaly here is associated with massive graphitic mineralization in northeast Bavaria, Germany (geologic context shown in (d)), close to the KTB borehole. Reproduced from Stoll J, Bigalke J, and Grabner EW (1995) Electrochemical modelling of self-potential anomalies. *Surveys in Geophysics* 16: 107–120.

(typically positive) countercharge migrates down the temperature gradient. For the diffusion of ions in response to a temperature gradient ∇T, eqn [1] can be written as (e.g., Revil, 1999)

$$\mathbf{j} = -\sigma_s \nabla V_{ed} + \sigma_s C_{te} \nabla T \qquad [24]$$

where C_{te} is the thermoelectric coupling coefficient (in V K^{-1}). For a porous medium saturated with a dilute, monovalent, symmetrical electrolyte, the coupling coefficient can be written as (Leinov et al., 2010)

$$C_{te} = -(T_+ - T_-)\frac{k_B}{e}\ln C + \frac{T_+}{F}\left(s^0_+ - \frac{Q^*_+}{T}\right) - \frac{T_-}{F}\left(s^0_- - \frac{Q^*_-}{T}\right) \qquad [25]$$

where $s^0_\pm$ denotes the partial molal entropy (J mol^{-1} K^{-1}) and $Q^*_\pm$ denotes the ionic heat of transport (J mol^{-1}), of the cations and anions, respectively. If the solid surfaces are not electrically charged, the SP is solely due to the thermal liquid-junction or diffusion potential, in which case the macroscopic Hittorf numbers $T_\pm$ are equivalent to the microscopic values $t_\pm$ and the thermoelectric coupling coefficient is given by

$$C^d_{te} = -(t_+ - t_-)\frac{k_B}{e}\ln C + \frac{t_+}{F}\left(s^0_+ - \frac{Q^*_+}{T}\right) - \frac{t_-}{F}\left(s^0_- - \frac{Q^*_-}{T}\right) \qquad [26]$$

Conversely, in the limit that all co-ions are excluded from the pore space (i.e., a perfect ion-selective membrane), the SP is solely due to the thermal membrane or exclusion potential and, assuming negatively charged mineral surfaces, the macroscopic Hittorf number $T_+ = 1$, and the thermoelectric coupling coefficient, is given by

$$C^e_{te} = -\frac{k_B}{e}\ln C + \frac{1}{F}\left(s^0_+ - \frac{Q^*_+}{T}\right) \qquad [27]$$

The sign of eqn [27] is reversed if the porous medium is positively charged. Assuming NaCl electrolyte (so $t_{Cl}/t_{Na} \approx 1.49$) and using values of $s^0_\pm$ and $Q^*_\pm$ from Lide (2009) and Agar et al. (1989), respectively, limiting values for the coupling coefficient are shown in **Figure 8**. Although it is recognized that thermoelectric processes may contribute to SP anomalies in geothermal areas (e.g., Corwin and Hoover, 1979; Darnet et al., 2004; Fitterman and Corwin, 1982), oil reservoirs (e.g., Gulamali et al., 2011; Jackson et al., 2012b), and underground fires (e.g., Corwin and Hoover, 1979; Dorfman et al., 1977; Revil et al., 2013), few experimental measurements of exclusion–diffusion potentials arising from temperature gradients in natural porous media have been reported. This is in marked contrast to the numerous published experimental measurements of streaming potential, covering a

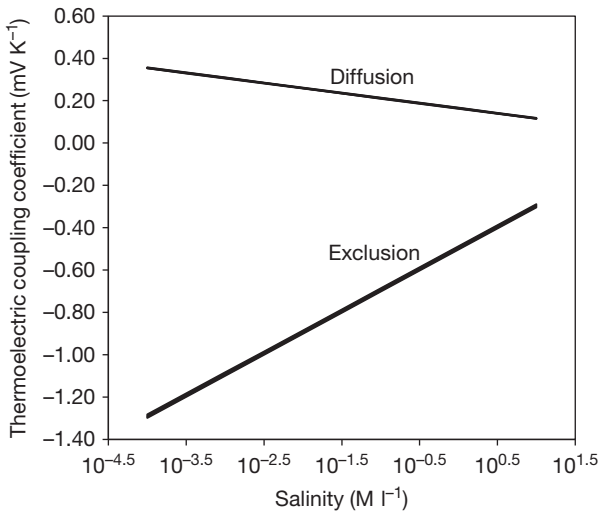


Figure 8 Limiting values of the thermoelectric coupling coefficient C_{te}, for the exclusion and diffusion dominated cases in a dilute, symmetrical, monovalent electrolyte, calculated using eqns [26] and [27] over the temperature range 15–55 °C (288–328 K). The effect of temperature is negligible, so the curves coincide. In all cases, $t_{Na}=0.39$ and the relative mobility is assumed constant and independent of temperature and salinity, which is reasonable over the range shown. Other parameters used in the calculation are taken from Lide (2009) and Agar et al. (1989). Modified from Leinov E, Vinogradov J, and Jackson MD (2010) Salinity dependence of the thermoelectric coupling coefficient in brine-saturated sandstones. *Geophysical Research Letters* 37: L23308. http://dx.doi.org/10.1029/2010GL045379.

wide range of natural samples of varying mineralogy and texture, saturated with electrolytes of varying composition and ionic strength (see, e.g., Section 1.2.6.7, and data in Vinogradov et al., 2010). Reported values of C_{te} vary from -1.36 to $0.25\ \text{mV K}^{-1}$ (see Leinov et al., 2010 for a review). Similar to exclusion–diffusion potentials of electrochemical origin (see **Section 11.09.2.2.1**), the value of C_{te} is dictated by the mobility contrast between ionic species, the thickness of the electric double layer relative to the radius of the water-occupied pores, and the surface charge of the porous medium (e.g., Leinov et al., 2010; Revil, 1999). If the double layer is thin compared to the pore radius, the ionic mobility contrast is large, or the surface charge on the porous media is small, then the diffusion potential dominates, and vice versa. Values of C_{te} lie between the end-members of pure diffusion and pure exclusion.

11.09.2.4 Telluric Currents, Stray Currents, and Sources of Noise in SP Surveys

The charge polarization mechanisms discussed so far all arise from earth processes, such as groundwater flow, concentration, temperature, or redox potential gradients, that may be the specific target of SP surveys or monitoring or that may contribute to an unwanted 'background potential' or 'noise.' For example, the well-known topographic potential, in which the SP becomes increasingly negative with elevation (the so-called negative summit effect), has been attributed to streaming potentials arising from groundwater flow and may be an unwanted background potential in, for example, a mineral exploration survey (e.g., Corry, 1985; Sato and Mooney, 1960; Telford et al., 1990). In other cases, such as in aquifer characterization, the streaming potential may be the signal of interest (e.g., Darnet et al., 2003; Jardani et al., 2009). However, there are a number of other naturally occurring electric current sources that are invariably undesirable in SP surveys, although these may also be the target of other survey types. Telluric currents are induced in the subsurface by variations in the Earth's magnetic field (e.g., Corwin, 1990; Telford et al., 1990). The main telluric current loops are very large (covering millions of square kilometers), are fixed with respect to the Sun, and rotate alternately clockwise and counterclockwise (Telford et al., 1990). Periodic and transient fluctuations in the associated electric fields, typically with periods of seconds to hours, can be correlated with diurnal variations in the Earth's magnetic field. Superimposed on this long-period telluric field are apparently random fluctuations with much higher frequencies (up to 100 kHz), related to electric disturbances in the ionosphere, such as those caused by electric storms. Consequently, the telluric record consists of erratic variations in electric potential, with amplitudes typically varying from μV to 100s mV and periods varying from μs to hours, which are recorded in SP surveys (Corwin, 1990). In areas of high resistivity or during electric storms, telluric currents may be associated with electric potential gradients of order $10–100\ \text{mV km}^{-1}$. These spatial and temporal variations in potential must be separated from the SP source(s) of interest.

Stray currents arise from anthropogenic sources including power lines and grounded electric machinery and are common in built-up areas. Stray currents from power lines typically have a frequency of 50–60 Hz and are easily eliminated by band-pass filters. Lower-frequency stray currents may have periods ranging from μs to hours and amplitudes of $100–1000\ \text{mV km}^{-1}$; these have a diverse range of sources, including electric railway lines, industrial plant, and domestic residences (e.g., Corwin, 1990; Nyquist and Corry, 2002). Identifying these sources may be essential to allow interpretation of measured data. One method for avoiding stray currents is to acquire data when the source is not operating. Other stray currents may be associated with corrosion of buried metal. The potentials associated with these may be large in magnitude (10–1000 mV) but generally have very long periods and may remain constant over several years. Vertical well casings often exhibit negative anomalies at the top, similar to ore bodies (e.g., see **Section 11.09.2.2.2**). Buried, unprotected pipelines and grounded elevated pipelines typically exhibit alternating positive and negative potentials. The long period of these potentials means that they are often simple to remove from survey data, as discussed in the next section.

11.09.3 Data Acquisition

Many applications of the SP method focus on surface and borehole SP surveys, in which data are acquired over a relatively short interval using traveling electrodes (e.g., Corry, 1985; Corwin, 1989, 1990; Sato and Mooney, 1960; Telford et al., 1990). However, there is increasing interest in long-term SP monitoring using repeat surveys, permanently installed electrodes, or a combination of both (e.g., Ernstson and

Scherer, 1986; Perrier and Pant, 2005; Trique et al., 2002). Borehole SP surveying is relatively routine, and numerous commercial companies offer this service using 'off-the-shelf' logging equipment (local or international providers can usually be identified via the Internet); in contrast, borehole monitoring projects, and surface surveying or monitoring projects (i.e., those that do not use a borehole logging tool), commonly utilize equipment assembled by the surveyor. Although numerous papers report the results of surface SP surveys, relatively few describe in detail the methods used to acquire the data; the interested reader is referred to papers by Corwin (Corwin, 1989, 1990) and a field SP survey manual by Corry et al. (1983). The basic equipment required for SP data acquisition in most situations is simple: (i) two or more nonpolarizing electrodes (discussed further in the next section), (ii) a high-resistance voltmeter to measure DC electric potential (internal resistance >1 GΩ), and (iii) insulated connecting wires. SP data acquisition is affected by error and noise; following Corwin (1990), error is defined here as the 'irreproducible component of a given SP reading associated with the data acquisition process,' while noise is a 'potential generated by a source that is not of interest for the purposes of the investigation.'

11.09.3.1 Electrode Types

A common feature of SP data acquisition, whether at the surface or in boreholes, or in field or laboratory experiments, is the use of so-called nonpolarizing electrodes, which typically comprise a metal (such as Cu, Ag, or Pb) immersed in a solution of a salt of the metal (such as $CuSO_4$); alternatively, if the metal salt has low solubility, the metal is coated with the metal salt and immersed in a conductive solution of a common anion (e.g., Ag–AgCl or Pb–$PbCl_2$ immersed in a solution of NaCl or KCl). A wide variety of electrodes have been designed and tested, which meet these basic requirements (e.g., Perrier et al., 1997). In many implementations, the metal electrode and electrolyte solution are located within a sealed container, with a low-permeability membrane (such as ceramic or wood) saturated with the same electrolyte allowing electric contact with the ground while minimizing leakage of the electrolyte (e.g., **Figure 9**). The ionic strength of the electrolyte depends upon the electrode type and application (e.g., monitoring in soil vs. water); a gelling agent may also be added to further reduce leakage of the electrolyte, or the electrolyte may be held within the pore space of a solid medium such as kaolinite or plaster. In some cases, the solution may be natural (e.g., seawater).

The advantage of nonpolarizing electrodes over the simple steel spikes often employed in resistivity surveys is that electrode polarization (i.e., the potential measured at a given time between an electrode pair in the absence of an external electric field) and drift (i.e., the time variation of the electrode polarization) are much smaller: typically less than a few mV compared to the several hundred mV across a pair of steel spikes. The main reason for this is that (ideally) polarization at the metal–electrolyte interface in each electrode is very similar in magnitude, approximately cancels across an (identical) electrode pair, and varies only slowly in response to slow variations in the electrolyte concentration. In such electrodes, also termed

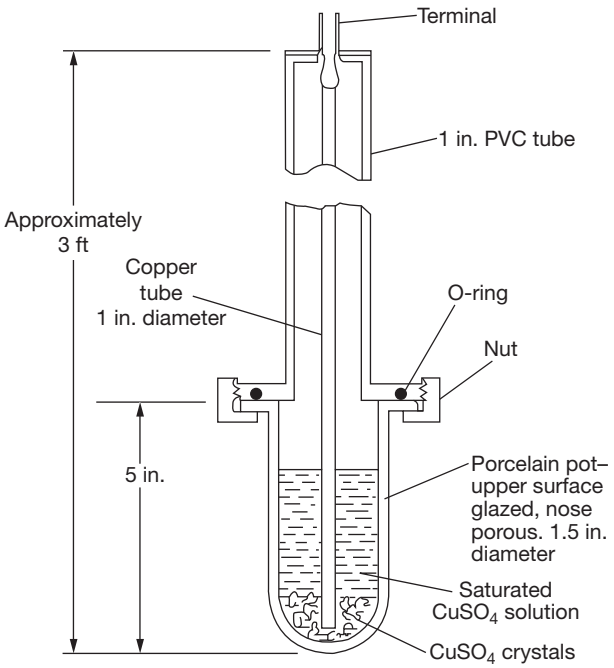


Figure 9 Example electrode designs for surface SP surveys. Reproduced from Telford WM, Geldart LP, and Sheriff RE (1990) *Applied Geophysics*, 2nd edn. Cambridge: Cambridge University Press.

'liquid-junction' or 'reference' electrodes, the 'effective' electrode polarization is the diffusion (liquid-junction) potential across the contact between the electrolyte solution and the adjacent saturated media, which has a much smaller magnitude and drift than the contact potential of a steel electrode. Note that the electrodes are not genuinely nonpolarizing, but the term is often used because of their observed behavior. When installed in water, the most significant environmental control on the effective electrode polarization is temperature; when installed in porous media such as soil, the most significant environmental controls are temperature and local moisture content (e.g., Corwin, 1990; Perrier et al., 1997).

Comparative studies of nonpolarizing electrodes for geophysical applications show that no one electrode type consistently exhibits lower polarization and/or drift; performance depends on the measurement conditions (e.g., Perrier et al., 1997; Petiau and Dupis, 1980). Ag–AgCl, Cu–$CuSO_4$, and Pb–$PbCl_2$ electrodes are the most widely used, with the metal electrode and electrolyte contained within a ceramic or plastic casing and a ceramic or wood membrane providing the electric connection (e.g., Corry et al., 1983; Corwin, 1990; Jackson et al., 2012a; Maineult et al., 2008; Perrier and Pant, 2005; Petiau, 2000; Vinogradov et al., 2010). Such electrodes are available from commercial suppliers (e.g., Cu–$CuSO_4$ and Ag–AgCl electrodes suitable for SP measurements in the field are manufactured for corrosion monitoring, while Ag–AgCl reference electrodes may be suitable for laboratory measurements); they are also relatively simple to manufacture in-house (e.g., the Pb–$PbCl_2$ electrodes of Petiau, 2000). Even using these electrodes, polarization and drift may be significant sources of error in SP surveys.

11.09.3.2 Laboratory SP Measurements

Laboratory measurements of SP in geologic porous media have been reported by a number of authors (see Section 1.2.6.7), often with the aim of determining the values of the cross coupling coefficients in eqns [3], [12], [23], and [24]. A key requirement of such experiments is to eliminate SP sources, including electrode polarization, that are not of interest. The experimental setup should be designed to ensure this. It may be useful to enclose the sample and measuring equipment within a Faraday cage to reduce the contribution of stray currents. The use of nonpolarizing electrodes is essential and, with these precautions, error and noise levels may be as low as a few μV (e.g., Jouniaux and Pozzi, 1995; Maineult et al., 2005; Vinogradov et al., 2010).

Streaming potential measurements should be conducted at uniform temperature, with negligible variations in electrolyte concentration and, no metallic components in contact with the electrolyte except for the electrodes. A typical setup to measure the streaming potential coupling coefficient comprises a confining system for the saturated sample (a simple column containing unconsolidated porous material or a pressure vessel holding a cylindrical plug of consolidated porous material), a method to impose a pressure differential across the sample so electrolyte flows through the pore space (a pump or a reservoir that can be elevated to increase the inlet head), two or more pressure transducers to measure the pressure difference across the sample, and two or more nonpolarizing electrodes to measure the voltage difference. Ideally, the temperature and electric conductivity or composition of the influent and effluent electrolyte should also be measured to ensure these remain constant. It may be necessary to isolate the electrolyte and sample from the atmosphere. If possible, the flow direction of the electrolyte should be reversed, to ensure the voltage response is symmetrical (i.e., electrode polarization and drift are small).

If flow through the sample is effectively 1-D, such that the pressure and voltage gradients are aligned, then the streaming potential coupling coefficient can be determined by measuring the voltage difference as a function of the pressure difference across the sample; at steady state and assuming no other SP sources, a linear regression through the data should yield a good fit ($R^2 > 0.9$) and the gradient of the regression gives the streaming potential coupling coefficient (e.g., eqn [4]; see also Section 1.2.6.4). Modifications to this simple experimental procedure may be required if the streaming potentials are very small (e.g., when the electrolyte is very saline; see Vinogradov et al., 2010), to acquire data at elevated temperature (e.g., Reppert and Morgan, 2003) or to investigate the effect of partial saturation (e.g., Guichet et al., 2003; Vinogradov and Jackson, 2011). Other flow or sample geometries may be preferred, depending upon the objective(s) of the experiment.

Measurements of SP of electrochemical origin have focused on (i) exclusion–diffusion potentials induced by concentration gradients, in the absence of pressure or temperature gradients (e.g., Ortiz et al., 1973; Revil et al., 2005a,b; Sen, 1989; Smits, 1968; Thomas, 1976) and (ii) SP signals associated with gradients in redox potential and buried electronic conductors (e.g., Castermant et al., 2008; Rittgers et al., 2013). The challenge in the former experiments is to eliminate the liquid-junction potentials that arise at the electrode–electrolyte interface; if the electrodes are immersed in electrolytes of differing concentration, then these 'electrode potentials' are different at each electrode. This is the case in experiments where the SP is measured across a sample in electric contact with electrolyte reservoirs of contrasting concentration. In this geometry, the measured SP can be directly related to the concentration difference across the sample but only after eliminating concentration-dependent electrode effects (e.g., Revil et al., 2005a,b). One approach, suitable for simple electrolytes, is to conduct a separate experiment, in which the two electrolytes are placed in electric contact (e.g., in a column with the denser electrolyte at the base), and the SP across the resulting liquid-junction is measured using the same electrodes. The liquid-junction potential for the given electrolyte pair is then calculated using eqn [16]; the electrode potential for this particular electrolyte–electrode pair is given by the difference between the calculated liquid-junction potential and the measured value (Leinov and Jackson, 2014). Another approach is to measure the SP using an experimental method in which the electrodes are located away from concentration gradients; see, for example, Ikard et al. (2012), who injected a saline tracer into a laboratory sandbox and monitored the SP response using electrodes at the sand surface. This approach eliminates concentration-dependent electrode effects, but the relationship between concentration gradient and SP is more challenging to determine. Similar sandbox experiments were used by Castermant et al. (2008) and Rittgers et al. (2013) to measure the SP response to the redox potential gradients associated with corrosion of an iron bar embedded in water-saturated sand, with the top exposed to atmosphere.

The requirement to eliminate electrode effects in thermoelectric potential measurements is similar to that in exclusion–diffusion potential measurements: If the electrodes are immersed in electrolytes of differing temperature, which is the case in experiments where thermoelectric potentials are measured across an insulated sample in contact with electrolyte reservoirs of contrasting temperature, then the electrode potentials are different. In this geometry, the thermoelectric potential can be directly related to the temperature difference across the sample, and electrode effects can be eliminated using a separate experiment similar to the approach described in the preceding text (e.g., Leinov et al., 2010; Leinov and Jackson, 2014). Alternatively, the SP can be measured in a geometry where the electrodes remain at constant temperature; see, for example, Dorfman et al. (1977), who inserted an electric heater into a large block of limestone and monitored the SP response using surface electrodes.

11.09.3.3 Borehole SP Measurements

Borehole measurements of SP are acquired using one of the oldest logging tools, which is, in principal, of very simple design: A single downhole electrode (often Pb) is connected to a reference electrode (or earth) at the surface, via a high-resistance voltmeter or galvanometer for the measurement of DC electric potential, and an adjustable DC voltage source that allows the measured potential to be offset to yield a baseline

signal of zero (e.g., Hearst et al., 2000; Jordan and Campbell, 1986). Occasionally, two downhole electrodes may be used in mineral exploration surveys, to yield an estimate of the SP gradient (e.g., Telford et al., 1990). In practice, the circuit may be rather more complicated, including inductors and capacitors to filter the signal, and be specific to the logging tool or company (e.g., Hearst et al., 2000; Jordan and Campbell, 1986). Current flow between the electrode and adjacent saturated rock occurs through the fluid(s) within the borehole, which must be electrically conductive; consequently, borehole SP logs can only be obtained in open holes (i.e., prior to the installation of casing) and in holes drilled with water-based mud.

Only changes in the measured SP from the chosen baseline are interpreted (see **Section 11.09.4.1**); the absolute value of measured potential is meaningless, because it includes the (unknown) electrode potentials (e.g., Hearst et al., 2000; Jordan and Campbell, 1986). These include chemical and thermal effects, because the surface and borehole electrodes are usually in contact with groundwater or drilling fluids of different chemical composition and temperature. A good reference electrode is necessary, which may be a spike driven into the ground close to, or remote from, the well head, or an electrode at shallow depth within the borehole (e.g., Hearst et al., 2000; Jordan and Campbell, 1986). A remote reference may help to reduce the effects of stray currents associated with activity at the well head, such as the use of electric generators, vehicles, and other industrial plant. Offshore measurements may be difficult to obtain because of the absence of a good earth on a rig and the amount of electric noise; the rig itself is rarely a suitable earth, because of the electric potential associated with the cathodic protection system. This may vary significantly with wave action, as the sacrificial ingots (typically of Mg) attached to the rig legs are exposed and submerged (see Wallace, 1968, for an entertaining account of the problems associated with offshore SP logging).

11.09.3.4 Surface SP Measurements

Surface SP measurements may be acquired (i) as part of a short-term surveying campaign, using one or more traveling electrodes, or (ii) during long-term monitoring, using traveling electrodes, a permanently installed electrode array, or a combination of these. In all cases, it is essential to ensure that there is a good electric connection between the electrode(s) and the ground and that variations in near-surface conditions, such as soil moisture content, are minimized (unless characterizing or monitoring these variations is of interest) (e.g., Corry et al., 1983; Corwin, 1990).

11.09.3.4.1 Electrode installation

It is usually necessary to dig holes at the SP survey stations to ensure a good electric connection with the ground (e.g., Corry et al., 1983). Ideally, the depth of the holes should be sufficient to (i) reach soil or rock that remains water-saturated regardless of the surface conditions, so an electric connection to deeper layers is maintained, and (ii) avoid streaming potentials associated with water flow in the near-surface layer affecting the measurements. To avoid introducing changes in moisture content, survey stations should be backfilled between measurements and their location flagged (Corwin, 1990).

Measurements of electrode contact resistance can be used to confirm that there is a good electric connection between the electrode and the ground, that there are no breaks in the electrode cable, and that the contact resistance is not so large that an appreciable loading of the measurement circuit may occur. Note that measurements of contact resistance cause current to flow through the electrodes and should be acquired after the SP measurement at a given station (Corwin, 1990). Some surveyors prefer not to water holes to decrease the contact resistance, as this can lead to nonuniform variations between survey stations (e.g., Corry et al., 1983); however, for very dry conditions, or if soil moisture varies considerably, watering may be necessary (e.g., Corwin, 1990). Several hours are required after watering prior to obtaining an SP measurement, to avoid recording transient streaming and electrochemical potentials. Even without watering, electrode readings may change rapidly within the first few seconds of installation, before reaching an approximately constant value; in some cases, the reading may continue to drift significantly after installation for several minutes or hours. Independent measurements and observations of factors such as soil moisture, soil type, temperature, weather conditions, vegetation, subsurface and outcrop geology, elevation, and telluric currents at each station during each SP measurement are very useful when interpreting the data (Corry et al., 1983; Corwin, 1990; Telford et al., 1990).

A different approach may be preferred for permanent electrode installations. An electrode configuration suitable for both permanent monitoring and repeat surveying, but requiring considerable effort to implement if there are numerous survey stations, has been suggested by Perrier and Pant (2005). A permanent monitoring electrode is installed at the base of a hole (of order 2 m in depth, but this may vary in response to local conditions), in a large (*c.* 40 l) plastic bucket filled with saturated or kaolinite mud (e.g., **Figure 10**). A plastic tube, also filled with saturated mud, connects the bucket to the surface, allowing measurements with a traveling electrode, inserted into the tube, to be obtained at the same survey station. The bucket reduces stray streaming potentials caused by local groundwater percolation and protects the electrode from biological activity and associated stray electrochemical potentials; the saturated mud ensures that the local moisture content remains approximately uniform and ensures a good electric contact with the ground. The use of coaxial cables, with the outer sheath connected to earth and the inner cable carrying the signal, can significantly reduce noise (Jackson et al., 2012a).

11.09.3.4.2 Base station/reference electrode installation

SP surveys, and monitoring arrays, record electric potentials relative to a reference electrode; this reference should preferably be installed at a base station outside of the survey or monitoring area, away from the SP source(s) of interest, but preferably in conditions that are similar to the monitoring arrays (e.g., if these are installed in dry, nonvegetated areas, the reference should not be installed in a wet, vegetated area). The reference electrode should be installed in the same configuration as the surveying or monitoring electrode(s) (e.g., Corwin, 1989, 1990; Perrier and Pant, 2005).

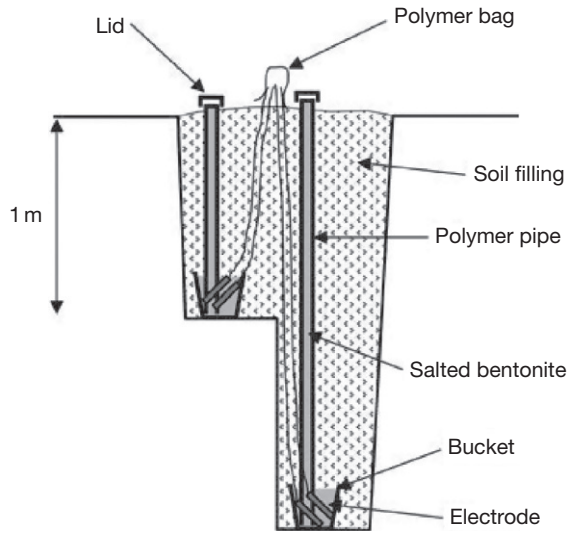


Figure 10 Example electrode installation for surface SP measurements. Reproduced from Perrier F and Pant SR (2005) Noise reduction in long-term self-potential monitoring with travelling electrode referencing. *Pure and Applied Geophysics* 162: 165–179.

11.09.3.4.3 Survey configurations

SP survey configurations can take a variety of forms, which have a number of advantages and disadvantages, in terms of the equipment and effort required and the errors introduced during data acquisition. In all cases, the negative terminal of the voltmeter should be connected to the reference electrode and the positive terminal to the measurement electrode (e.g., Corwin, 1989, 1990; Telford et al., 1990). It is generally desirable to lay out survey lines in loops, so electrode drift can be quantified at loop closure points (e.g., Corwin, 1990).

In the simplest approach (often called the 'fixed-base' or 'total field' method), a single traveling electrode measures the potential at each survey station relative to a reference electrode at the base station. The advantages of this approach are that (i) only two electrodes are required, (ii) measurement errors do not accumulate (in contrast to the gradient or leapfrog methods discussed next), and (iii) errors introduced by electrode polarization and time-varying stray currents can be estimated by measuring the potential difference between the reference and traveling electrode at the base station immediately before electrode installation and at the end of the survey (Corwin, 1990). If the reference electrode is installed for long periods (*c*. hours), drift may be estimated by measuring the reference potential against a second electrode or by returning the traveling electrode to obtain a measurement at the base station. If the potential difference between the reference and traveling electrode, measured at the base station at time t, is δV_{t-r}, then the 'true' potential at measurement station n at time t is related to the 'measured' potential by

$$V_n^{\text{true}}(t) = V_n^{\text{meas}}(t) - \delta V_{t-r}(t) \qquad [28]$$

A plot of δV_{t-r} against t and appropriate curve fit to these data (requiring a minimum of two points for a linear regression; higher-order regressions require additional measurements of δV) can be used to determine the appropriate correction at intermediate times. Multiple electrode holes may be occupied at each reference station, to help quantify the effect of local variations in soil conditions.

The disadvantage of the fixed based method is that a very long cable is required (capable of reaching the most distant survey station), which may not be logistically possible and increases the possibility of cable damage. Nevertheless, where possible, the fixed-base method is generally preferred (Corwin, 1990).

An alternative approach, which requires a much shorter cable length, is the gradient method (also termed the dipole, leapfrog, or fixed electrode configuration; Corwin, 1990). The method again requires two electrodes, but measurements are made by moving the dipole along the survey line and measuring the potential between the electrodes. The trailing electrode occupies the previous station of the leading electrode. If the station spacing is small, then the measurements essentially yield the potential gradient dV/ds at the station midpoint (where s is the station spacing), which can be plotted for interpretation; alternatively, successive values of dV can be added to yield the SP at a given station relative to the starting or base station (Telford et al., 1990). Reversal of the leading and trailing electrodes between stations (the so-called leapfrog technique) can help to reduce the cumulative error caused by electrode polarization.

The advantage of the gradient method is the relatively short cable length, which may be essential in some surveys. The disadvantage is that measurement errors are cumulative and can dominate the later measurements. The cumulative errors can be reduced by adding interconnecting survey lines at numerous tie-in points (i.e., survey stations shared by multiple loops) and distributing tie-in loop closure errors (i.e., the difference in voltage measured at common survey stations in different survey loops) among all readings within a given loop. Yet this can lead to a situation where the data to be interpreted differ significantly depending on the number and location of interconnecting survey lines; in this case, the measurement errors are still unacceptably large. The use of the gradient method should be restricted to those cases where it is unavoidable for logistic reasons (Corwin, 1990). A compromise may be possible, in which a number of base stations are established for the fixed based method, thus reducing the total cable length required, with multiple tie points to help reduce the errors between base stations (Corwin, 1990).

A multielectrode survey configuration, or a long-term SP monitoring network, utilizes an electrode at each survey station, with the potential measured with respect to a common reference at the base station, or as a series of dipoles, preferably recorded using a multichannel data acquisition system (e.g., Corwin, 1990; Perrier and Pant, 2005; Trique et al., 2002). Such a setup allows repeat measurements that can be used to determine the time variation of SP sources of interest; moreover, cumulative error is reduced by the use of a single base station. However, the initial costs of a multielectrode array are considerably higher; furthermore, errors introduced by electrode polarization and time-varying stray currents may be difficult to identify. Errors in the permanent array measurements can be monitored using a traveling electrode, with measurements obtained at the survey stations via the installation shown in **Figure 10**. Assuming the traveling electrode T yields the 'true' potential difference across a given permanent dipole

P1 − P2 measured at time t, the correction δV_{P1-P2} to the potential difference measured across the permanent electrodes $V_{P1} - V_{P2}$ is given by

$$\delta V_{P1-P2}(t) = V_{P1}(t) - V_{P2}(t) - V_{T1}(t) + V_{T2}(t) \qquad [29]$$

Similar to the fixed-base survey method, appropriate corrections for permanent electrode measurements acquired at intermediate times can be determined using a curve fit to a plot of δV against t.

Once the survey type and configuration has been determined, the location, spacing, and number of survey stations must be identified. This will be driven by the objectives of the survey or monitoring program and, ideally, some prior knowledge of the SP anomaly of interest. A preliminary survey may be required if the area is unknown; for example, an initial leapfrog survey with relatively wide dipole spacing may be used to delimit gross variations in SP, prior to identifying a suitable base station and the desired areal coverage for a more detailed fixed-base or multielectrode survey or monitoring program.

11.09.4 Data Interpretation, Modeling, and Inversion

Interpretation of borehole SP logs, and surface SP surveys, is often largely qualitative or semiquantitative; for example, borehole logs are typically used to identify permeable bed boundaries by eye, with perhaps some quantitative estimates of formation water salinity, while surface SP surveys are used to identify anomalies that may be associated with buried ore bodies. The shift toward more quantitative interpretation approaches started with the works of Fitterman (1978, 1979, 1981) and Sill (1983). Fitterman (1978, 1979) developed integral equations to derive the electric and magnetic fields associated with current sources in the Earth, while Sill (1983) developed a framework to solve the differential form of the governing equations using numerical methods. Various authors have developed these initial studies to formulate the forward problem in terms of integral equations relating the source distribution to the measured SP (e.g., Birch, 1993, 1998; Fournier, 1989; Revil et al., 2002, 2004) or solve the forward problem using numerical methods (e.g., Saunders et al., 2008, 2012; Sheffer and Oldenburg, 2007). In all cases, it is essential to identify the SP sources that are acting to produce the observed response; if these are unknown, ambiguous or incorrectly specified, all interpretations will be in error.

11.09.4.1 Interpretation of Borehole SP Logs

Numerous textbooks report the basic principles of SP log interpretation, typically for oil and gas applications, so only a brief summary is provided here (see, e.g., Hearst et al., 2000; Jordan and Campbell, 1986; Telford et al., 1990). The key link between application and interpretation is the SP source acting to produce the observed response; in most SP logging applications, the SP is assumed to be electrochemical in origin, arising from the contrast in chemical composition between the drilling fluids (mud filtrate) and the natural formation water (e.g., Hearst et al., 2000; Jordan and Campbell, 1986). Borehole SP logs drilled in other settings, such as for mineral exploration, may have SP sources that are different or additional to the electrochemical sources considered here (e.g., Goldie, 2002; Stoll et al., 1995).

A borehole SP log yields a continuous record of the potential difference between a traveling downhole electrode and a fixed reference electrode located either at surface or at a shallower level within the borehole, as a function of depth (see **Section 11.09.3.3**). Prior to interpretation, the temperature dependence of the electrode response must be corrected for; typically, this is assumed to be linear and is removed either by using a known temperature coefficient for the electrode or by applying a linear regression to the measured data to determine the temperature coefficient (e.g., Hearst et al., 2000; Jordan and Campbell, 1986). The SP signal is then interpreted, assuming it arises from (i) diffusion potentials resulting from concentration differences between mud filtrate and formation water and (ii) exclusion potentials across mudstones or shales. The (constant) SP recorded when the traveling electrode is adjacent to a thick shale bed is termed the 'shale baseline'; conversely, the SP recorded when the traveling electrode is adjacent to a thick, clay-free sandstone bed is termed the 'sandline'; the difference between the two is termed the static SP (SSP). The recorded SP is typically adjusted to yield a shale baseline of zero, prior to interpretation (e.g., Hearst et al., 2000; Jordan and Campbell, 1986). If the sandstone beds contain significant clay minerals or thin shale beds, the full SSP may not develop and the SP difference is termed the pseudostatic SP (PSP). If the mud filtrate is less saline than the formation water, then the SSP and PSP are negative, and vice versa.

The magnitude of the SSP primarily depends upon the salinity contrast between the mud filtrate and the formation water. Ideally, the shale baseline corresponds to the exclusion (membrane) potential given by eqn [16], and the sandline corresponds to the diffusion potential given by eqn [17], so the SSP (i.e., the difference between these potentials) is given by (e.g., Hearst et al., 2000; Jordan and Campbell, 1986)

$$V_{SSP} = V_d - V_e = -\frac{2k_B T}{e}(1 - t_+)\ln\left(\frac{\sigma_w}{\sigma_{mf}}\right) \qquad [30]$$

where salinity C has been replaced by conductivity σ. If $t_+ = t_{Na} = 0.39$ and using $\ln x = 2.303 \log x$, then eqn [30] becomes

$$V_{SSP} = V_d - V_e = -0.24T\log\left(\frac{\sigma_w}{\sigma_{mf}}\right) \qquad [31]$$

where the SSP is expressed in mV. If the mud filtrate conductivity σ_{mf} is known and the SSP can be identified, then eqn [31] (sometimes termed the 'standard SP model') can be used to estimate the formation water conductivity σ_w (e.g., Hearst et al., 2000; Jordan and Campbell, 1986). Note that eqn [31] is valid only if the mud filtrate and formation water have relatively low salinity, the dominant salt species is NaCl, and the rock is fully water-saturated. If the salinity is high or there are other salt species present, then the solution activity must be used instead of conductivity; often, conductivity is replaced by an equivalent value (or its inverse, an equivalent resistivity) that is defined to be proportional (or inversely proportional)

to activity (e.g., Hearst et al., 2000; Jordan and Campbell, 1986).

In practice, the SSP may not be observed, because (i) the presence of clay minerals in sandstones causes some charge exclusion from the pore space, (ii) co-charge is not entirely excluded from mudstones, and (iii) the SP anomaly in permeable beds, relative to the shale baseline, is reduced by invasion of mud filtrate, restricted bed thickness, increased borehole diameter, the presence of hydrocarbons, and/or low formation conductivity (e.g., Doll, 1948; see **Figure 11**). A number of studies have proposed corrections for these effects (e.g., Segesman, 1962; Worthington and Meldau, 1958). The contribution of streaming potentials arising from the flow of mud filtrate in the adjacent formation may also need to be removed by estimating values for the pressure gradient and coupling coefficient, although these are usually neglected (e.g., Wyllie, 1951). Qualitative interpretation of SP logs typically allows (i) identification of the boundaries between mudstones and porous beds at the inflection point of the SP curve, (ii) correlation of beds across multiple wells based on similar SP characteristics, and (iii) estimate of the mud content of sandstones (i.e., how 'clean' they are) based on trends in SP within beds; a relative decrease in the SP response toward the zero shale baseline may indicate an increase in clay content (e.g., **Figure 12**). A number of papers have developed quantitative approaches to SP log interpretation (see, e.g., Woodruff et al., 2010 and references therein), typically in conjunction with other borehole logging tools.

11.09.4.2 Interpretation of Field SP Surveys and Monitoring

Surface SP surveys may be used instead of, or in addition to, borehole logging to determine the SP at a specific time or over a short time interval; repeat surveying, or long-term monitoring, may determine the time variation in the SP. The end results of a surface SP survey are a set of SP profiles or a contour map of equipotentials. Borehole data may also be available to constrain the depth variation of the SP. Interpretation of SP surveys and monitoring campaigns has, until recently, largely been qualitative or semiquantitative, particularly when surveys

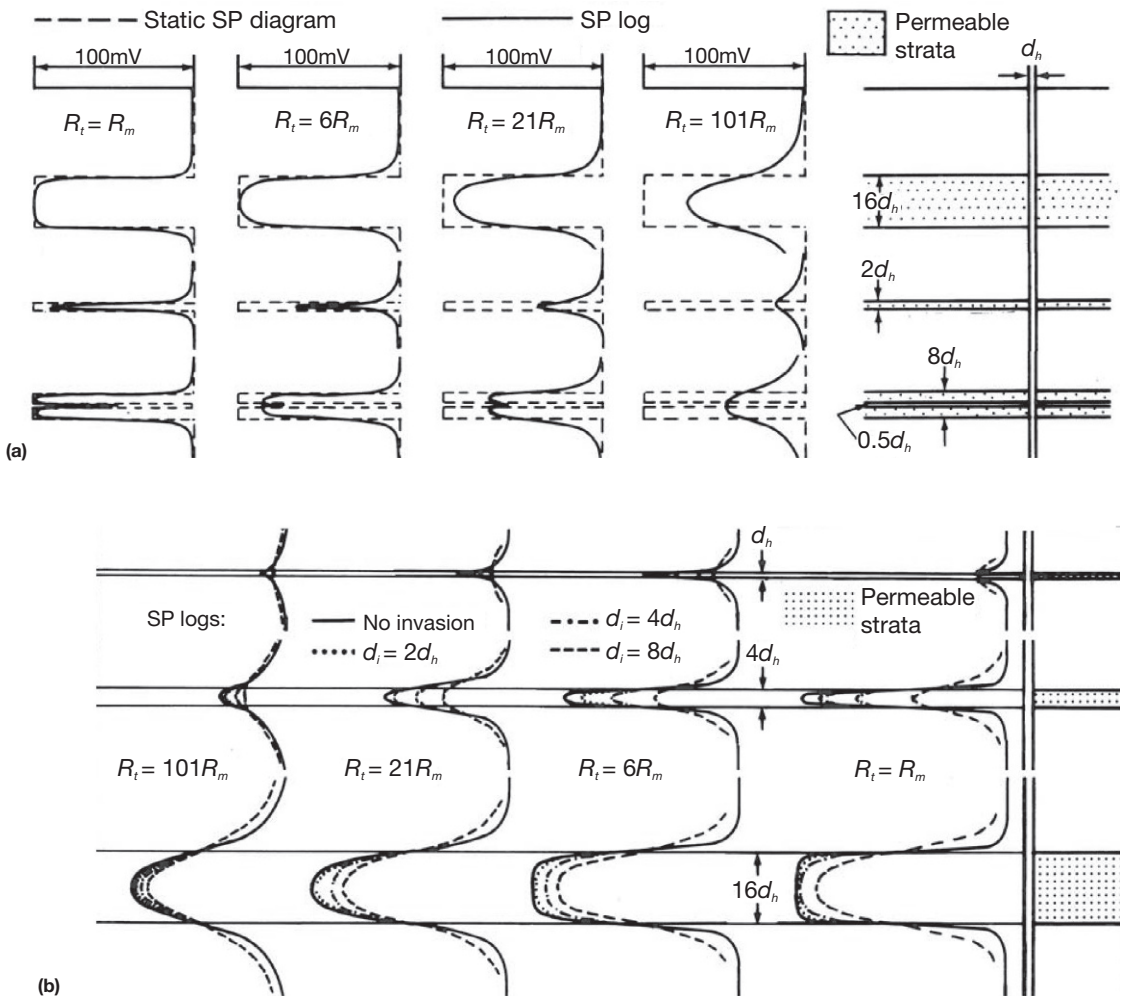


Figure 11 Example SP logs showing the reduced deflection from the shale baseline (a) as bed thickness decreases and the ratio R_t/R_m increases, where R_m denotes the mid filtrate resistivity and R_t denotes the total formation resistivity, and (b) mud filtrate invasion increases. Reproduced from Jordan JR and Campbell FL (1986) *Well Logging II – Electric and Acoustic Logging, Vol. 10: SPE Monograph*, 159 pp.

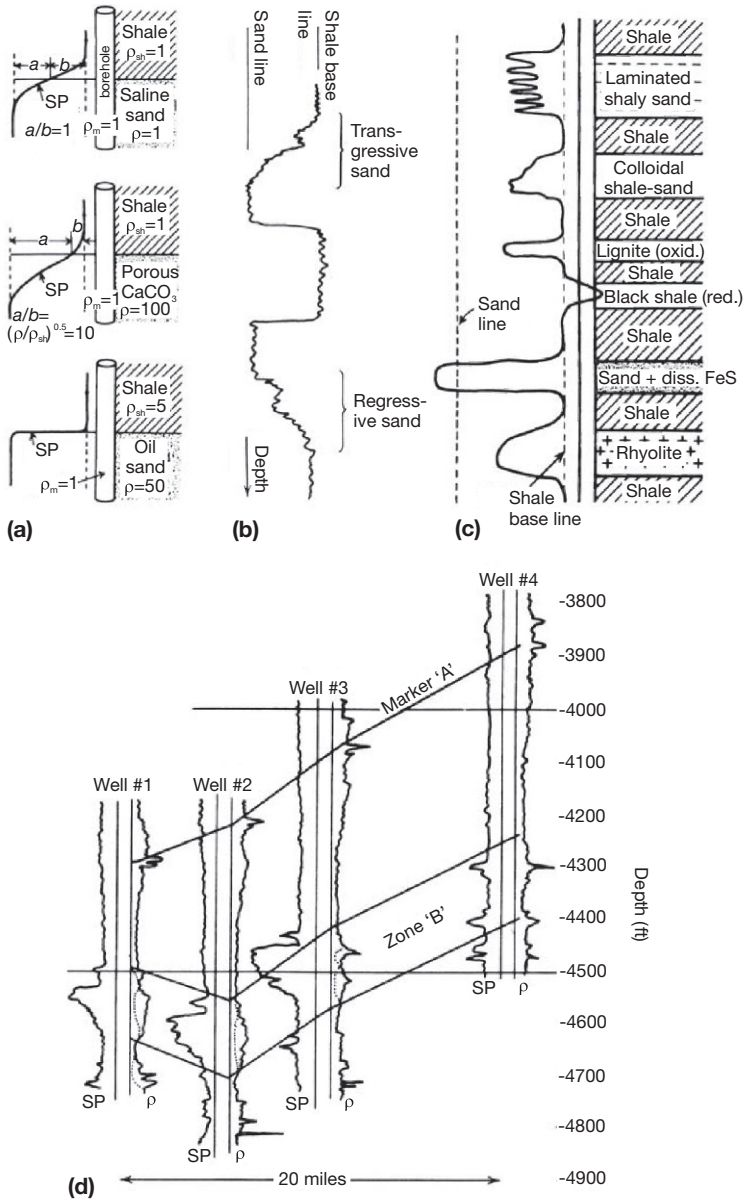


Figure 12 Example quantitative interpretations of SP logs. (a) Identification of permeable bed boundaries from the inflection in the SP curve. (b) Qualitative estimation of mud content in sandstone: as the SP deflection moves toward the sand line, the mudstone content decreases downward in the downward-coarsening transgressive sands and decreases upward in the upward-coarsening regressive sandstones. (c) Relationship with lithology: note the effect of the thin bedding in the laminated sand-shale bed, the asymmetrical curves for colloidal sand-shale and for rhyolite, the positive anomaly for a reducing bed, and the large anomaly associated with pyrite. (d) Correlation of beds across different wells. Modified from Telford WM, Geldart LP, and Sheriff RE (1990) *Applied Geophysics*, 2nd edn. Cambridge: Cambridge University Press.

have been undertaken for mineral exploration, because the relationship between ore bodies and the associated SP anomaly is relatively poorly understood (e.g., Telford et al., 1990; see also **Section 11.09.2.2.2**). The shape of an anomaly and its lateral and vertical extent are indicated by the SP profiles or contour map; the depth of the source may be obtained from borehole data or, crudely, the shape and amplitude of the anomaly: if $x_{1/2}$ is the total width of the profile at half the maximum amplitude, then the depth to the top of the source may be of order $x_{1/2}/2$, although there are numerous exceptions to this (Telford et al., 1990). An approximate estimate of the attitude of the source may also be obtained from the symmetry of the anomaly (or lack thereof): the steep slope and positive tail of the anomaly are typically on the down-dip side. As with all potential field problems, the SP source inversion problem is highly nonunique, and the challenge of interpreting the data is exacerbated by the relatively sparse measurement locations and presence of noise.

Such crude 'rules of thumb' are informed by potential distributions calculated for polarized bodies of varying shape (e.g., dipoles, spheres, and ellipsoids; **Figure 13**), which have been used to provide semiquantitative interpretations of

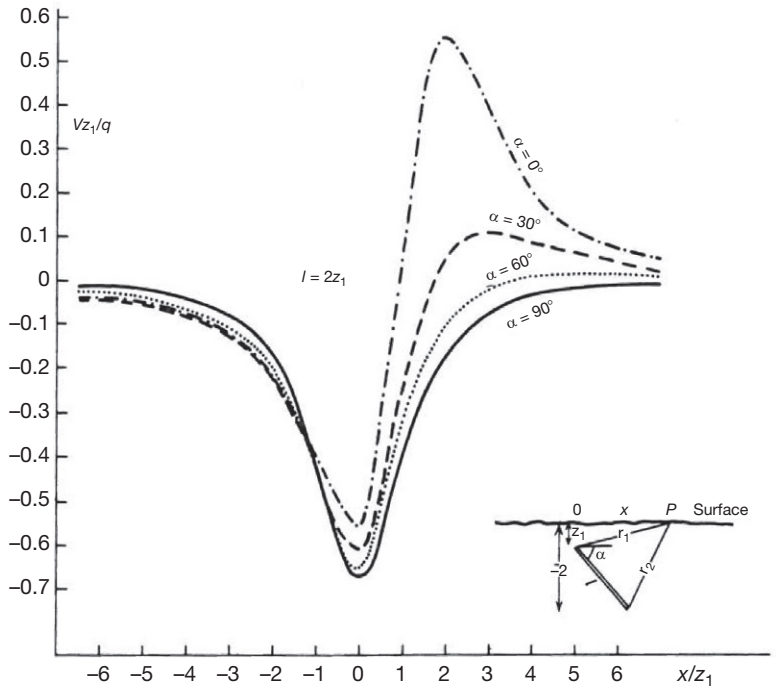


Figure 13 SP profiles calculated over a buried polarized rod. Reproduced from Telford WM, Geldart LP, and Sheriff RE (1990) *Applied Geophysics*, 2nd edn. Cambridge: Cambridge University Press.

measured SP anomalies in a number of studies (e.g., Meiser, 1962; Murty and Haricharan, 1985; Paul, 1965; Rao and Babu, 1984). However, these techniques provide little information about the nature of the SP current sources. Moreover, heterogeneity in the conductivity structure of the over- and underburden can significantly affect the SP measured at surface; for example, sulfide bodies have been observed to produce large (*c.* 200 mV) and zero surface SP anomalies for sand and clay overburden, respectively (Telford et al., 1990). Prior to interpreting an anomaly, it is often necessary to remove regional effects or other SP variations not associated with the source of interest; this is typically done by inspection, based on observations of surface conditions, subsurface geology and the nature of the overburden, possible anthropogenic sources such as buried pipelines or nearby machinery, and knowledge of other SP sources in the area. The problem with such an empirical approach to data analysis is that it is specific to the interpreter and cannot be easily reproduced.

Recent studies, particularly those with environmental or hydrogeological applications, have begun to develop quantitative methods to interpret SP surveys, usually in conjunction with other geophysical data, particularly resistivity surveys. For example, Jardani et al. (2007) proposed a method to invert SP measurements to determine the 3-D distribution of Darcy velocity in the shallow subsurface, assuming (i) the measured SP is dominated by the streaming potential associated with groundwater flow, (ii) there is a simple relationship between the excess charge density transported by the flow and the permeability of the porous media, and (iii) there is a known electric conductivity structure. Their approach was extended to jointly invert SP and temperature data by Jardani et al. (2009). Minsley et al. (2007b) used a 3-D SP inversion algorithm to determine an electric current source model, accounting for the 3-D resistivity structure derived from a resistivity survey at the same field location, and showed that the sources and sinks of electric current, determined from the inversion, can be related to zones of high and low redox potential. However, the match between observed and modeled contaminant concentration was of variable quality. Naudet et al. (2003) and Arora et al. (2007) used piezometric data to remove the streaming potential component from SP survey data and correlated the residual electrochemical signal with direct measurements of redox potential. Mendonça (2008) applied the model of Stoll et al. (1995) in both forward and inverse modeling of SP signals associated with inert ore bodies. Castermant et al. (2008) combined the approaches of Arora et al. (2007) and Mendonça (2008) to interpret SP signals associated with a corroding iron bar. These and similar studies are based upon a model of the underlying SP current sources, as will be discussed in the next section.

11.09.4.3 SP Forward Modeling Concepts

Forward and inverse numerical modeling of the 3-D distribution of SP current sources (i.e., the second term on the right-hand side of eqn [1]) lie at the heart of quantitative approaches to interpret SP data, whether measured at the surface or downhole (e.g., Ahmed et al., 2013; Gulamali et al., 2011; Ishido and Pritchett, 1999; Jardani et al., 2007, 2008; Mendonça, 2008; Minsley et al., 2007a; Sailhac and Marquis, 2001; Saunders et al., 2008, 2012; Sheffer and Oldenburg, 2007; Sill, 1983). The SP current sources must be related to the coupled thermodynamic potentials; as discussed in **Section 11.09.2**, these may be gradients in pressure,

composition, redox potential, temperature or a combination of all of these. Following Sill (1983), the general equation for coupled flows can be written as

$$\mathbf{\Gamma}_j = \sum_k L_{jk}\mathbf{X}_k \quad [32]$$

where the fluxes $\mathbf{\Gamma}_j$ (charge, matter, heat, etc.) are related to the various forces $\mathbf{X}_k$ (gradients of electric potential, pressure, composition, temperature, etc.) through the coupling coefficients L_{jk}. When the flux is the electric current $\mathbf{j}$, eqn [32] is equivalent to eqn [1]. The SP problem is to model 'secondary' electric current flows and potentials (i.e., $j=2$), which are driven by some other primary flow ($j=1$); couplings between the primary flows are not of direct interest here. Typically, the effects of the secondary electric potentials on the primary flow are assumed to be small, so the primary flow equation can be decoupled from the secondary flow equation, yielding (Sill, 1983)

$$\mathbf{\Gamma}_1 = -L_{11}\nabla\psi \quad [33]$$

$$\mathbf{\Gamma}_2 = \mathbf{j} = -\sigma_s\nabla V - L_{21}\nabla\psi \quad [34]$$

where $\mathbf{\Gamma}_1$ is the primary flow (matter, heat, etc.), L_{11} is the primary conductivity (permeability, thermal conductivity, etc.), Ψ is the primary potential (pressure, temperature, etc.), and L_{21} is the appropriate cross coupling coefficient for SP current source i in eqn [1]. Specific equivalent expressions to eqn [34] for streaming, electrochemical, redox, and thermoelectric current sources are given by eqns [3], [12], [23], and [24], respectively; the L_{21} cross coupling coefficient is related to the coupling terms C_i (where i denotes the SP source) by

$$L_{21}^i = -\sigma_s C_i \quad [35]$$

The decoupled primary flow problem (eqn [33]) can be solved separately using the most convenient method available and the values obtained used in the solution of the electric flow problem (eqn [34]). The second term in eqn [34] is a 'convection' current driven by the gradient in primary potential, and the first term is the usual Ohmic conduction current driven by the gradient of the electric potential. Equation [34] can therefore be written as

$$\mathbf{j}_{\text{total}} = \mathbf{j}_{\text{cond}} + \mathbf{j}_{\text{conv}} \quad [36]$$

where

$$\mathbf{j}_{\text{cond}} = -\sigma_s\nabla V \quad [37]$$

$$\mathbf{j}_{\text{conv}} = -L_{21}\nabla\psi \quad [38]$$

If there are no external current sources or sinks, then the total current is divergence-free (eqn [2]) so

$$-\nabla\cdot\sigma_s\nabla V = \nabla\cdot(L_{21}\nabla\psi) = \nabla L_{21}\cdot\nabla\psi + L_{21}\nabla^2\psi \quad [39]$$

Equation [39] shows that there are sources (nonzero divergence) of conduction current wherever there are gradients in the cross coupling coefficient L_{21} parallel to the primary flow and wherever there are external or induced sources of the primary flow. The sources of the conduction current given by the right-hand side of eqn [39] can be used to determine the resultant SP (V). The $\nabla L_{21}\cdot\nabla\psi$ term in eqn [39] shows that gradients in the same direction produce conduction current sources (positive current sources), while gradients in the opposite direction produce conduction current sinks (negative current sources) (e.g., **Figure 14**). Substitution of eqn [33] into the $L_{21}\nabla^2\psi$ term results in two additional source terms

$$L_{21}\nabla^2\psi = \frac{-L_{21}\nabla\cdot\mathbf{\Gamma}_1}{L_{11}} + \frac{L_{21}\nabla L_{11}\cdot\mathbf{\Gamma}_1}{L_{11}^2} \quad [40]$$

The first of these is nonzero at externally imposed sources of the primary flow. If $\nabla\cdot\mathbf{\Gamma}_1 > 0$, a source for the primary flow, then this corresponds to a conduction current sink if $L_{21} > 0$ (i.e., $C_i < 0$, which is typically the case in rocks with negatively charged mineral surfaces) and vice versa (e.g., **Figure 14**). The second term is nonzero at boundaries where the primary flow conductivity changes. These boundaries are the locations of the secondary or induced sources for the primary flow.

Solution of the SP follows from solution of the primary flow equation, which is used to determine ψ within the domain of interest, plus any other properties (such as salinity and temperature) that affect the values of the primary conductivities L_{11} and coupling terms L_{21}. The first step in a numerical solution of eqns [33] and [34] is to use an appropriate program to solve the primary potential problem (fluid flow, heat flow, etc.). The second step consists of using the primary potential solution along with a model for the cross coupling coefficients to calculate the sources for the electric problem from eqn [34]. The final step uses the current sources along with an electric conduction model to determine the resultant electric potentials. A number of authors report numerical methods to solve the primary and secondary flow equations for a variety of applications (e.g., Gulamali et al., 2011; Mendonça, 2008; Saunders et al., 2008, 2012; Sheffer and Oldenburg, 2007; Sill, 1983). The boundary conditions used to solve the electric problem are typically a Neumann condition at the earth surface, such that $\mathbf{n}\cdot\sigma_r\nabla V = 0$ where $\mathbf{n}$ is the unit vector normal to the surface, and a Dirichlet condition within the Earth, such that $V \to 0$ as the distance from the current sources becomes large. The latter condition typically requires a large model domain. A reference value of SP should be specified at a fixed location within the model, which may mimic the location of a reference electrode in a field survey (e.g., Saunders et al., 2012).

11.09.4.4 Inversion of SP Measurements

With resistivity inversion, the target is the conductivity structure, which reproduces the potentials measured for a given current source distribution (e.g., Herwanger et al., 2004; Pain et al., 2003); with SP inversion, the target is the location, strength, and type of the current sources, which yield the measured data, within a conductivity structure that may also be unknown or uncertain. Inversion of SP data to date has assumed a known electric conductivity structure (e.g., Ahmed et al., 2013; Jardani et al., 2007, 2008; Mendonça, 2008; Minsley et al., 2007a; Patella, 1997). The SP inverse problem involves the calculation of the source terms throughout the domain from an incomplete set of measured potentials. Although the SP source inversion is a linear problem, there are two fundamental difficulties. The first is the large variation in sensitivity of the data to different model parameters; the second is the ill-posed nature of the problem with a nonunique solution, compounded by the limited set of

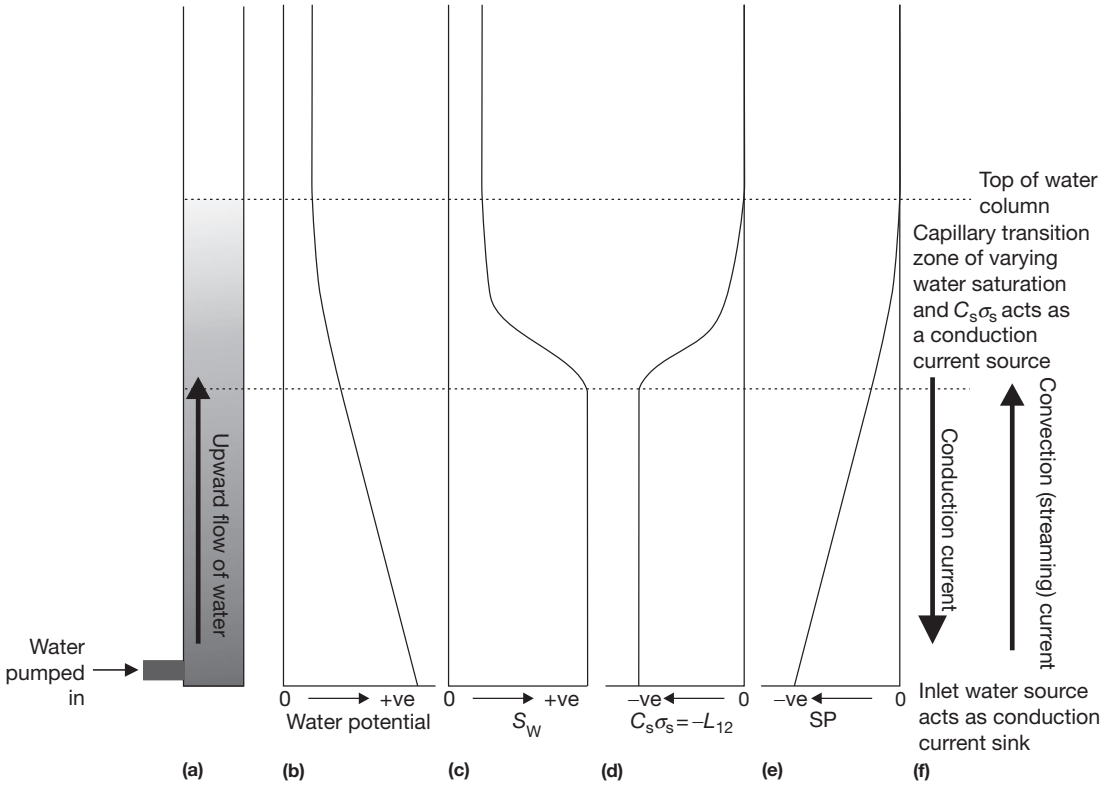


Figure 14 Illustration of relationship between gradients in the cross coupling coefficient L_{21} and conduction current sources given by eqn [39], using the example of an SP anomaly associated with water being pumped into a vertical, sand-filled column from the base (a). The water potential (b) decreases upward; the water saturation (c) and cross coupling coefficient $L_{21} = -\sigma_s C_s$ (d) also decrease upward owing to the saturation dependence of σ_s (given by Archie's second law) and C_s (given by eqn [11]). Note that C_s is negative, consistent with negatively charged mineral surfaces in the sand. The gradient in L_{21}, in the same direction as the water potential gradient (where $\nabla L_{21} \cdot \nabla \psi > 0$), acts as a conduction current (positive) source. The inlet source of water (where $\nabla \cdot \mathbf{\Gamma}_1 > 0$) acts as a conduction current sink (negative source), yielding an increasingly negative streaming potential with depth, measured relative to an electrode at the column outlet (e).

available measurements (e.g., Minsley et al., 2007a). The basic inversion problem, regardless of the modeled SP mechanisms, can be expressed in discretized form as (e.g. Minsley et al., 2007a)

$$\mathbf{K}^{-1}\mathbf{s} = V \qquad [41]$$

where $\mathbf{s}$ is the matrix of current sources at each spatial location in the discretized model (i.e., the discrete form of the right-hand side of eqn [39]) and $\mathbf{K}$ is a positive definite, symmetrical matrix that represents the discretized form of $\nabla \cdot \sigma_r \nabla$ including boundary conditions (left-hand side of eqn [39]). $\mathbf{K}$ includes the earth conductivity structure, so errors in the conductivity structure will clearly result in SP source estimate errors. The source matrix $\mathbf{s}$ could theoretically be calculated if V is known completely, but in practice, this is never the case. Instead, V is sampled at a limited subset of locations that are often restricted to the earth surface, as described in **Section 11.09.3**.

The nonuniqueness of the solution has been addressed in several ways. One approach is to select a class of source solutions that have a desired spatial compactness, reflecting the often spatially localized nature of SP current sources. This choice amounts to the addition of prior knowledge about the properties of the sources, which may depend on the problem of interest. Mendonça (2008) and Castermant et al. (2008) inverted SP measurements associated with ore bodies and a corroding iron bar, respectively, assuming the SP sources are located at the surface of the ore or metal body, consistent with the exchange current model of Stoll et al. (1995) described in **Section 11.09.2.2.2**. Similarly, Jardani et al. (2009) inverted SP measurements to identify the location of the water table, assuming that the SP sources coincide with the nonzero divergence of the streaming current density at the location of the water table. Minsley et al. (2007a) imposed spatial compactness at the regularization step and obtained solutions using increasingly compact models (e.g., **Figure 15**), but without assuming a priori the spatial location of the SP sources, and applied this approach to invert SP measurements associated with contaminant plumes (Minsley et al., 2007b) and aquifer pumping tests (Minsley et al., 2007a). Another approach to handle the nonuniqueness of the SP inversion problem incorporates Tikhonov regularization with a smoothness constraint based on the differential Laplacian operator and a prior model, which can be determined from a forward model of the problem of interest and a best guess of the required material properties (e.g., Ahmed et al., 2013; Jardani et al., 2006, 2007, 2008).

The inverse problem is solved using an optimization method to minimize the error between the simulated and observed data, subject to the constraints on the current sources

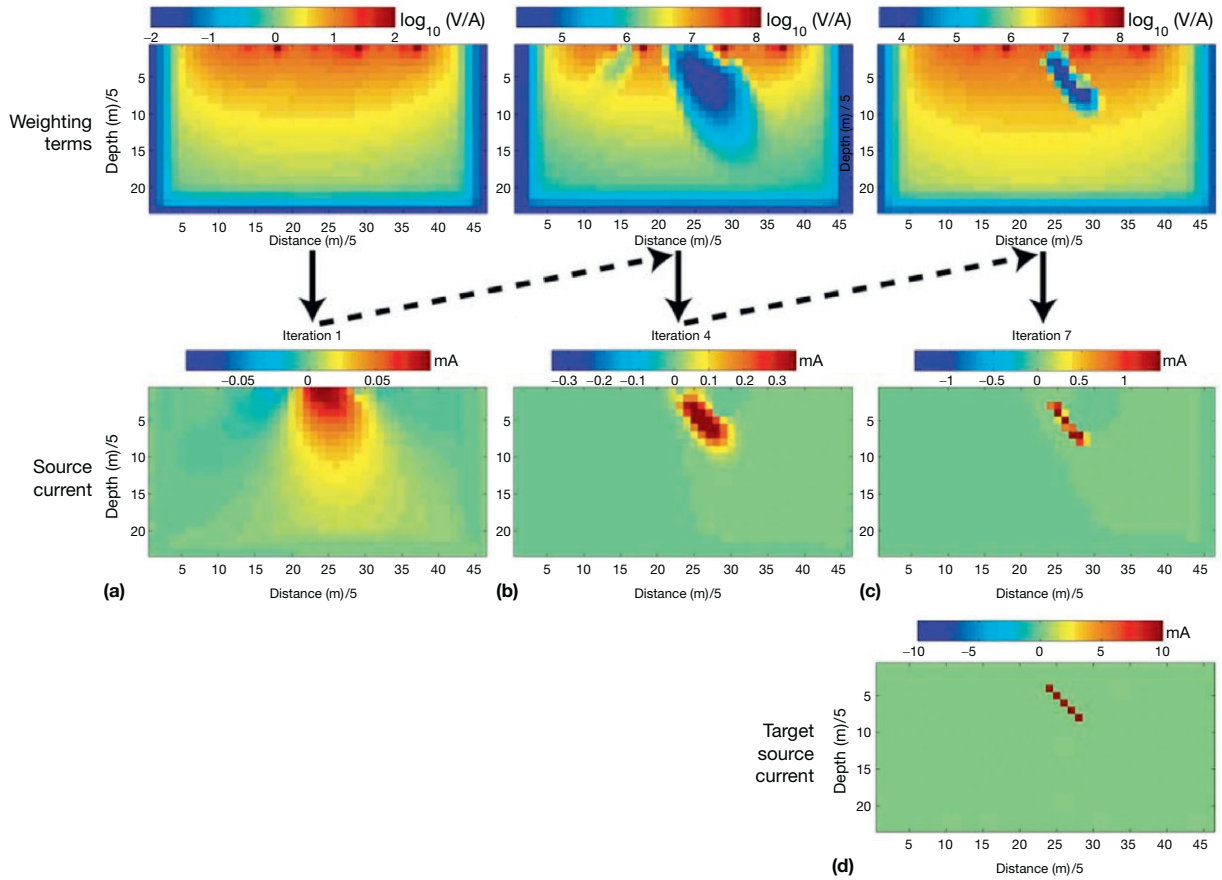


Figure 15 Impact of imposed source compactness on inversion of SP data from a synthetic current source. A 2-D section through a 3-D problem is shown, with (top row) the iterative weighting terms and (middle row) the source inversion results for iterations (a) 1, (b) 4, and (c) 7. Iteration 1 corresponds to the sensitivity weighted minimum norm solution. Note the different color scale for each iteration, where iteration 7 recovers approximately 20% of the true source amplitude shown in (d). Modified from Minsley BJ, Sogade J, and Morgan FD (2007) Three-dimensional source inversion of self-potential data. *Journal of Geophysical Research* 112: B02202. http://dx.doi.org/10.1029/2006JB004262.

described earlier. Optimization methods used for SP inversion so far include least squares (e.g., Jardani et al., 2006; Patella, 1997) and conjugate gradient (e.g., Minsley et al., 2007a,b). Regardless of the approach used, solution of the inverse problem yields a distribution of current sources that must be related to the coupled driving forces (gradients in pressure, temperature, etc.) using eqns [3], [12], [23], and [24]. This step is, therefore, specific to the problem of interest. Jardani and coworkers (Ahmed et al., 2013; Jardani et al., 2007, 2008) modeled source terms that were electrokinetic in origin and assumed that the streaming charge density Q_w can be related to the permeability k according to a relationship of the form

$$\log_{10} Q_w = h \log_{10} k + g \qquad [42]$$

with $h = -0.82$, $g = -9.2$ and k in m^2 (Ahmed et al., 2013). This information was then used to calculate the Darcy flow field. Mendonça (2008) and Castermant et al. (2008) modeled source terms that were electrochemical in origin and given by the exchange current model of Stoll et al. (1995) (eqn [22]), while Naudet et al. (2003) and Arora et al. (2007) assumed they were given by eqn [23]. Minsley et al. (2007b) assumed a simple linear relationship between the current source density and contaminant concentration.

11.09.5 Selected Case Studies

In this section, some case studies are briefly reviewed to demonstrate applications of SP surveying and monitoring. In some cases, SP data were acquired and interpreted in isolation; in others, they were acquired as part of a suite of geophysical survey data.

11.09.5.1 Mineral Exploration

SP surveys are commonly used in mineral exploration, and this case study by Corry (1985) summarizes the acquisition and qualitative interpretation of some 3000 line km of survey data over both known and prospective sulfide deposits in the United States (Corry, 1985). The data were acquired using the survey techniques outlined in Section 2.6.2.4: nonpolarizing electrodes (typically the Cu–$CuSO_4$ type), the fixed-base surveying method, shallow pits (*c.* 10–15 cm) for reference and measurement electrodes, and measurements every 100–200 m along a line. If these techniques are closely followed, Corry (1985) found that survey loops *c.* 20 km in circumference tied within ±25 mV. Loop errors can be further reduced, but the

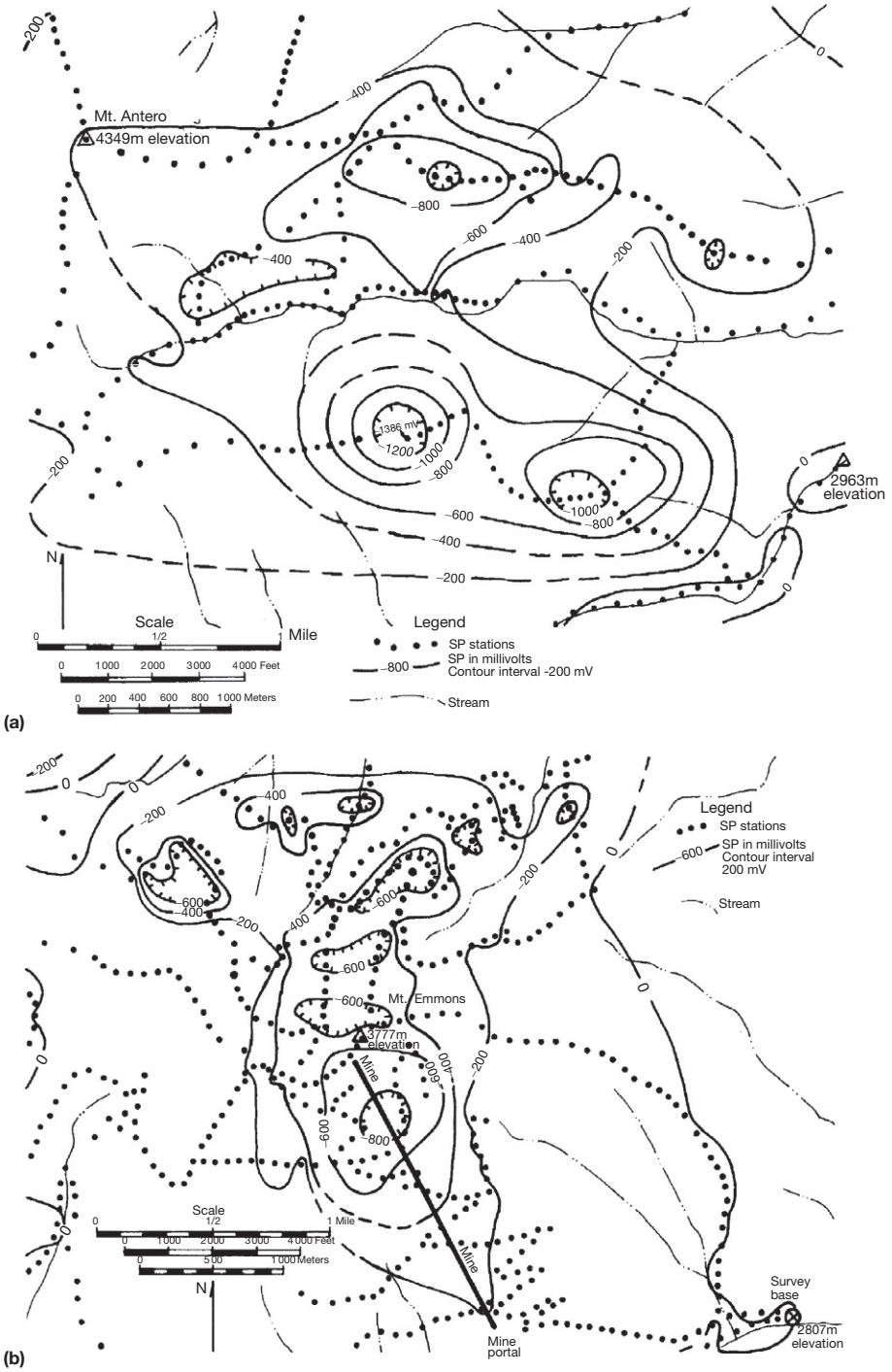


Figure 16 Negative surface SP anomalies associated with buried sulfide ore bodies in (a) the Sawatch Mountains, southwest of Buena Vista, and (b) Mt. Emmons near Crested Butte, CO. Reproduced from Corry CE (1985) Spontaneous polarization associated with porphyry sulfide mineralization. *Geophysics* 50: 1020–1034.

increased costs of such careful surveying are not usually justified in a mineral survey, where the amplitude of the anomalies commonly exceeds several hundred millivolts.

Figure 16 shows typical results from two separate surveys: negative SP anomalies of magnitude >200 mV are associated with sulfide systems in the Sawatch Mountains, southwest of Buena Vista, and Mt. Emmons, near Crested Butte, CO. No further quantitative interpretation of the data was made, beyond identifying the SP anomaly by eye and confirming the presence of sulfide deposits by direct sampling. Repeat surveys over several years in the area shown in Figure 16(b) revealed no significant variations, indicating that the SP anomalies are stable through time. Moreover, SP data obtained from the mine shown in Figure 16(b) indicate that negative

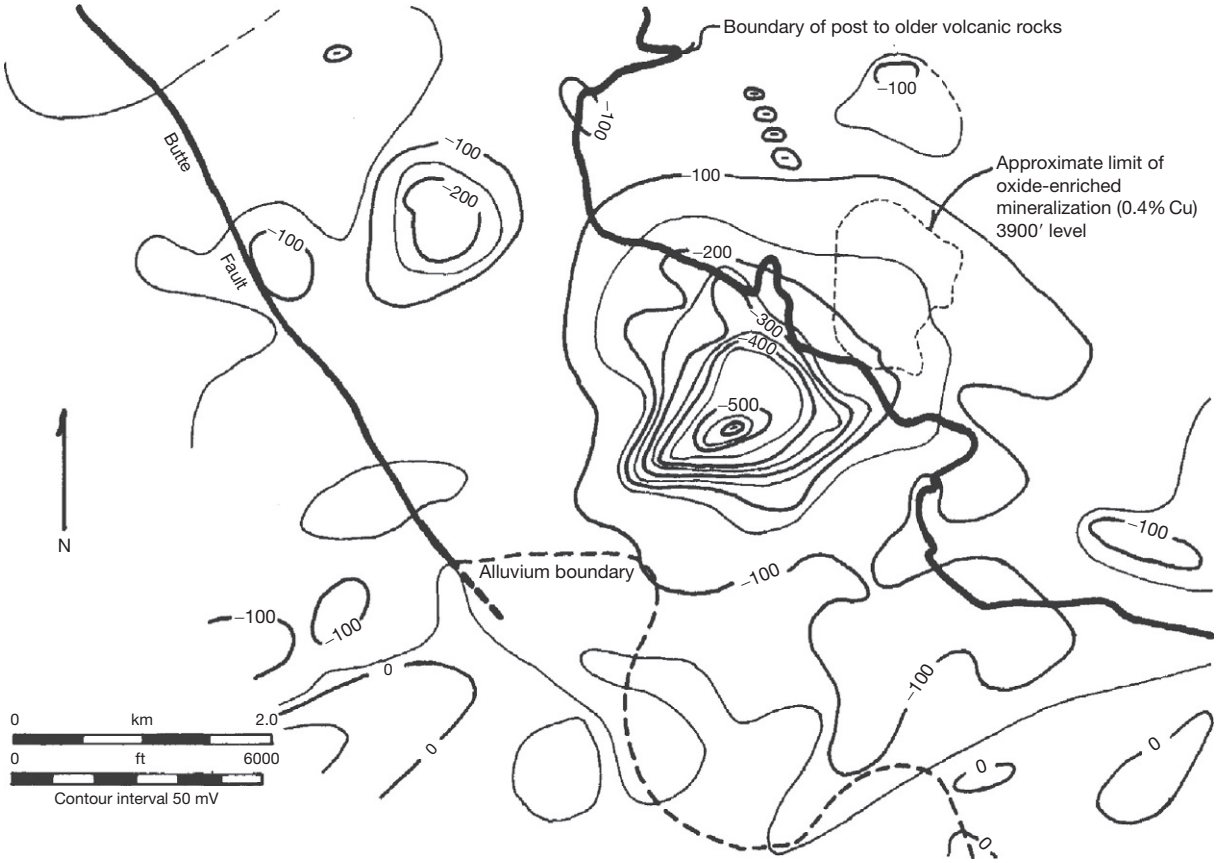


Figure 17 Surface SP anomaly above the Safford deposit, Arizona. A known, oxide-enriched ore body at approximately 204 m below ground surface has no direct associated SP anomaly; the prominent SP anomaly of amplitude −600 mV located *c.* 1.5 km southwest of the original deposit was found by later drilling to be associated with a deeper deposit at *c.* 550 m below ground surface. Reproduced from Corry CE (1985) Spontaneous polarization associated with porphyry sulfide mineralization. *Geophysics* 50: 1020–1034.

anomalies are also present at depth. However, sulfide deposits are not invariably associated with SP anomalies (e.g., **Figure 17**). In the Safford deposit, Arizona, a known, oxide-enriched ore body at approximately 200 m below ground surface has no direct associated SP anomaly; the prominent SP anomaly of amplitude −600 mV located *c.* 1.5 km southwest of the original deposit was found by later drilling to be associated with a deeper deposit at *c.* 550 m below ground surface. Corry (1985) argued that such ore bodies are associated with SP anomalies only if they have not oxidized through exposure to weathering, although the mechanism they proposed is inconsistent with the IEM model of Stoll et al. (1995) (see **Section 11.09.2.2.2**).

11.09.5.2 Environmental Site Surveying

Naudet et al. (2003, 2004) and Arora et al. (2007) reported the use of SP surveying to delimit the extent of a contaminant plume migrating from the Entressen landfill site, located in the south of France approximately 70 km northwest of the city of Marseille (**Figure 18**). The Entressen landfill is the largest open-air landfill in Europe and has operated since 1912. Municipal and domestic waste is stored directly on the ground surface with no impermeable bottom liner. The landfill extends over ~0.5 km^2 and reaches a height of 30 m above the regional land surface. Borehole data suggest that landfill leachates percolate to a shallow unconfined aquifer (1–8 m depth) formed by the paleoalluvial plain of the Durance river. The portion of the aquifer impacted by the landfill is composed of high-permeability quaternary alluvium a few meters thick (hydraulic conductivity of *c.* 8.5×10^{-3} m s^{-1}), underlain by a thick layer of low-permeability marl (hydraulic conductivity of *c.* 10^{-6} m s^{-1}). The average hydraulic gradient of the aquifer in the vicinity of the landfill is 5%, with a flow direction predominantly to the southwest, but locally modified by the presence of a few high-permeability paleochannels. The yearly fluctuation of the water level is about 1 m. The contaminated portion of the aquifer is characterized by high electric conductivity of the groundwater, approximately ten times higher than the reference measured (~1 km up-gradient of the landfill), and predominantly negative values of the redox potentials. The negative redox potentials indicate the presence of an anaerobic zone in the contaminated plume.

To aid in delimiting the lateral extent of the contaminant plume, an SP survey was conducted between September 2001 and March 2002, using an MX20 Metrix multimeter (*c.* 10^8 Ω internal impedance), one 500 m reel of wire, and two Pb–$PbCl_2$ Petiau electrodes (Petiau, 2000). The electrodes were embedded in *c.* 10 cm holes, filled with *c.* 50 g of salty, wet bentonite mud, to ensure a good electric contact with the dry and rocky ground. To reduce the influence of local heterogeneities, the SP signal was measured at five locations within

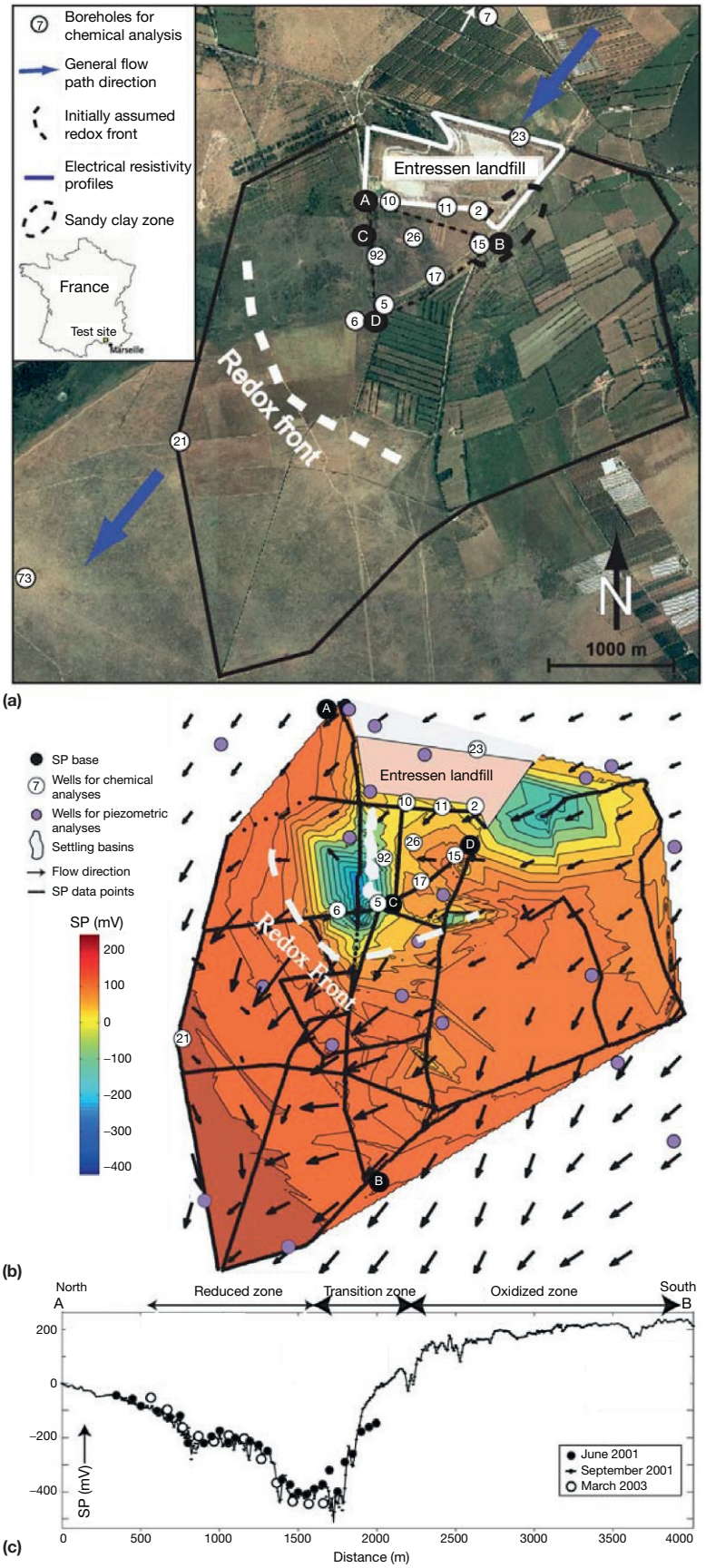


Figure 18 Contaminant plume and associated SP anomaly at the Entressen landfill site. (a) Location and associated monitoring of the down-gradient contaminant plume. The location of the redox front associated with the plume, estimated using borehole data, is also shown. (b) SP survey over the Entressen landfill site. (c) One SP survey line running approximately N–S along line A–B in (b), measured on three different dates. Modified from Naudet V, Revil A, Bottero J-Y, and Bégassat P (2003) Relationship between self-potential (SP) signals and redox conditions in contaminated groundwater. *Geophysical Research Letters* 30(21): 2091. http://dx.doi.org/10.1029/2003GL018096.

1 m^2 of each station, and the average used to represent the station value; these measurements had a standard deviation of 15–20 mV. A combination of the gradient and fixed-base survey techniques was used to reduce cumulative errors, and survey lines were interconnected at numerous tie-in points. Closure errors were redistributed over the base stations of the loop, yielding residual errors of *c.* 5 mV per base station. All SP data were finally referenced to a master base station located up-gradient of the landfill. Data were acquired with a 10 m sample spacing near the landfill and a 20 m spacing further down-gradient to the southwest. The resulting SP map, generated using triangular interpolation, is shown in **Figure 18(b)**. The SP survey showed good repeatability (**Figure 18(c)**).

In addition to the SP survey, thirteen electric resistivity tomography (ERT) profiles were obtained in the first km down-gradient of the landfill, within the triangle ABD in **Figure 18(a)**, using an ABEM multielectrode Terrameter (SAS 4000) equipment in the Wenner configuration and 64 stainless steel electrodes with a uniform spacing of 3 m in each profile. The ERT imaged (1) the vadose zone, the aquifer, and the marly aquitard, which have strongly contrasting resistivity, and (2) correlated conductivity with that measured in boreholes for the contaminated groundwater. Geochemical sampling and measurements performed between 1998 and 2000 demonstrate that the contaminant plume extends to a maximum of 4.6 km down-gradient of the landfill (to well #73) with an anaerobic zone located in the two first kilometers (to well #6) (Vilomet et al., 2001).

Naudet et al. (2003, 2004) recognized that the total SP anomaly observed in **Figure 18(b)** has two principal sources: firstly, a streaming potential component associated with groundwater flow and, secondly, a component associated with the gradient in redox potential associated with the plume. The streaming potential component was removed, initially by simply correlating the measured SP with hydraulic head at well locations outside of the plume where the redox contribution is negligible (e.g., **Figure 19(a)**) and later using the Bayesian method developed by Linde et al. (2007) and described in Arora et al. (2007). The residual SP anomaly was then correlated with E_h values measured in boreholes (**Figure 19(b)**). Finally, a map of E_h was developed (**Figure 19(c)**). Naudet et al. (2004) contoured the extent of the contaminant zone using the boundary of $E_h = 0$, yielding a higher-resolution map of the redox front than was possible using borehole data alone (cf. **Figure 18(b)**). Thus, SP data provide an inexpensive method for improving the spatial resolution of mapping of contaminant plumes. Note that, in **Figure 19(c)**, the negative redox potentials to the east of the landfill are likely associated with a clay zone, while the very large negative potential (*c.* −1 V) to the southwest of the landfill is likely associated with the cathodic protection system of a buried pipe.

11.09.5.3 Leak Detection from Dams

Several studies have suggested the use of SP surveys and monitoring to identify leaks and preferential flow paths through dams and embankments, and the study described here, reports a simple case where SP surveying was able to identify the flow path from a leaking dam and how this changed through time (Titov et al. 2000). The dam studied is located in Petergoph, a suburb of St. Petersburg, Russia, and separates a water storage pond, originally constructed to supply a fountain center similar to Versailles in France, from a canal and river (**Figure 20**). The dam is constructed of sand, and the subsurface consists of sand, gravel, and sandy clay. The water table below the dam is at a depth of *c.* 1 m. The survey reported by Titov et al. (2000) was conducted after the second of two leakage incidents damaged the dam structure during spring floods, including a spring forming on the river bank *c.* 100 m from the dam.

Resistivity and SP surveys were carried out in April and May 1998. A single resistivity profile was obtained approximately parallel to the dam (**Figure 20(a)**) and a fixed station SP survey was carried out in the area between the dam and the river bank (**Figure 20(b)** and **20(c)**), using Ag–AgCl electrodes and the same high-impedance receiver as the resistivity survey, with survey stations located on a 4×8 m grid. Data processing prior to interpretation was simple and consisted of correction of the measurement electrode drift using return measurements at the base station every 15 min. The resistivity survey results yielded a three-layer model of the subsurface, consisting of a shallow layer of unsaturated, high-resistivity sand, underlain by a layer of low resistivity saturated sand and a base layer of clay. The first SP survey, conducted in April, showed a strong anomaly associated with a locally thickened area of sand and oriented along a line running from the dam to the spring, with the negative end at the dam and the positive end at the spring (**Figure 20(b)**). Titov et al. (2000) interpreted the anomaly to be a streaming potential caused by water leaking from the pond and through the locally thickened layer of permeable sand in the shallow subsurface. The polarity of the anomaly is consistent with a negative streaming potential coupling coefficient (the negative potential is associated with the higher water head and vice versa, as in **Figure 14**). This in turn is consistent with a negative ζ, as typically observed in sands and sandstones (e.g., **Figure 2(b)**). Given the approximately linear nature of the SP anomaly, parallel to the flow direction, Titov et al. (2000) estimated a value of the streaming potential coupling coefficient and corresponding ζ using eqn [4] (see **Section 11.09.2.1**), which assumes 1-D flow so the net current density is zero. The resulting values of -2.5×10^{-7} V Pa^{-1} and −100 mV are reasonable for the conductivity of the pond water measured (*c.* 22 Ωm, corresponding to a salinity of *c.* 3×10^{-3} M; see **Figure 2**).

A second SP survey conducted in May showed a much less pronounced SP anomaly, suggesting much less water was percolating through the shallow subsurface; a significant reduction in spring water discharge was also observed (**Figure 20(c)**). Titov et al. (2000) suggested that thawing of temporarily frozen soils in and immediately below the dam provided an alternative route for water to leave the pond and discharge to the swampy area in between the dam and the river. SP measurements in this study were used to image temporal changes in flow paths in the shallow subsurface with a minimum of quantitative interpretation.

11.09.5.4 Hydrogeology

To contrast with the simple, qualitative interpretation of SP data presented in the previous case study, we finish with a case study in which inversion of surface SP measurements, in conjunction with a 3-D resistivity survey, was used to identify the

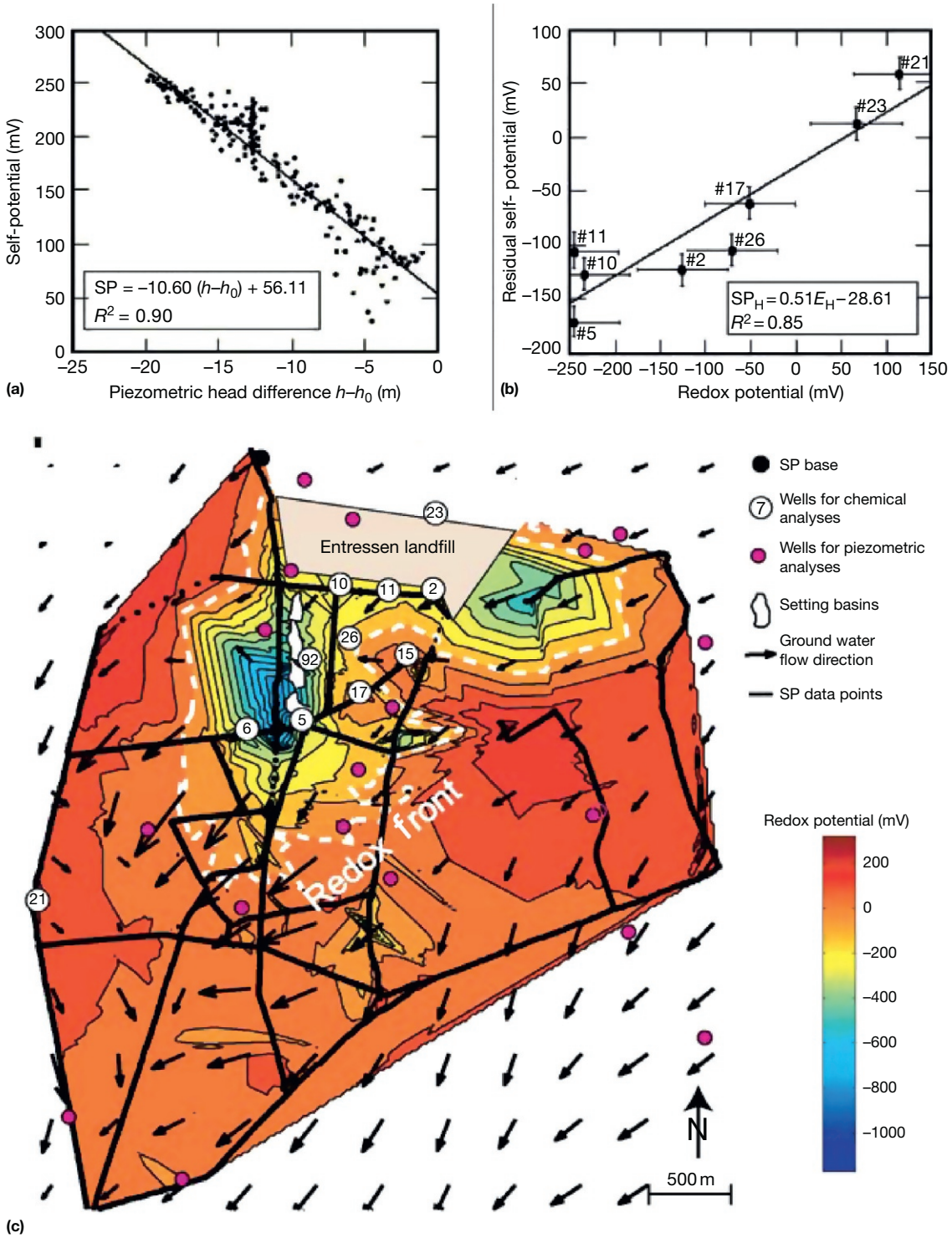


Figure 19 Interpretation of SP data for redox potential. (a) Correlation between SP and hydraulic head at borehole locations away from the contaminant plume, used to remove the streaming potential contribution to the measured SP. (b) Correlation between residual SP, after removing the streaming potential component, and redox potential E_h measured at boreholes. (c) Map of redox potential obtained from the residual SP and the correlation shown in (b). Modified from Naudet V, Revil A, Rizzo E, Bottero J-Y, and Bégassat P (2004) Groundwater redox conditions and conductivity in a contaminant plume from geoelectrical investigations. *Hydrology and Earth System Sciences* 8(1): 8–22.

current sources associated with a plume of hot groundwater rising to the ground surface in the geothermal field of Cerro Prieto, Baja California, Mexico (Jardani et al., 2008; **Figure 21(a)** and **21(b)**).

The Cerro Prieto geothermal field lies in a tensional zone between the Cerro Prieto and the Imperial transform faults. The heat source for the hydrothermal system is probably located to the east, near well NL-1 (**Figure 21(c)**) in an area

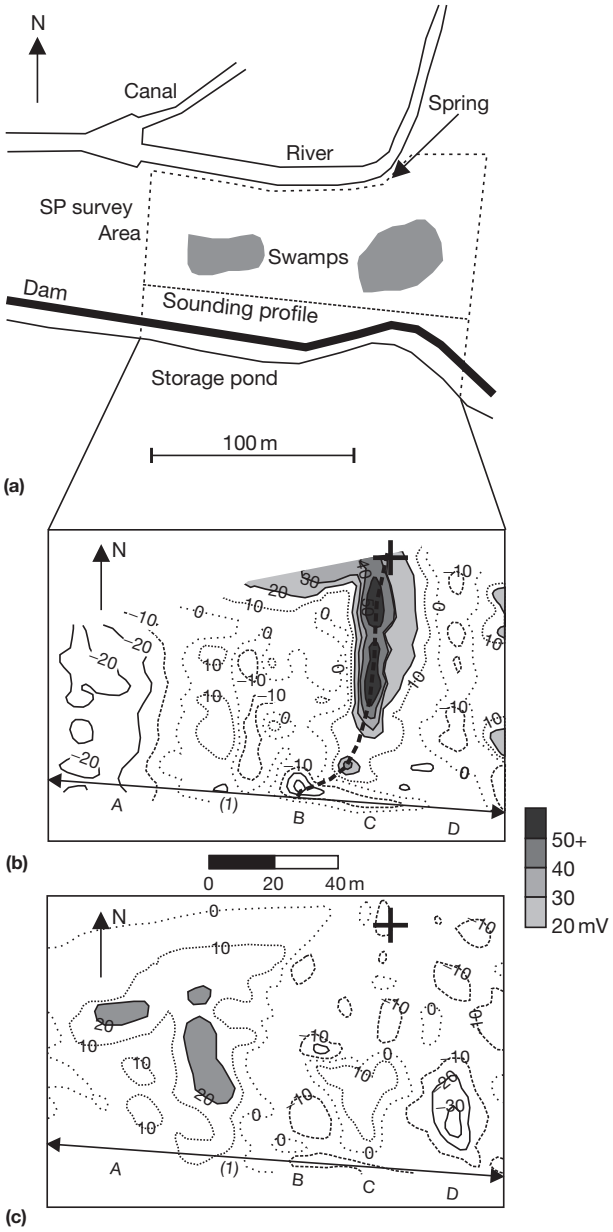


Figure 20 The part of the Petergoph fountain water supply system studied by Titov et al. (2000). Sketch map (a) and SP survey maps obtained in (b) April and (c) May. Line (1) denotes the resistivity sounding profile, letters (A–D) denote regions of higher subsurface sand thickness identified in resistivity data, the bold dashed line in (b) denotes the SP anomaly due to leakage from the dam in April, and the cross denotes the location of the spring being fed by the leak. Modified from Titov K, Loukhmanov V, and Potapov A (2000) Monitoring of water seepage from a reservoir using resistivity and self-polarization methods: Case history of the Petergoph fountain water supply system. *First Break* 18: 431–435.

where wells have drilled mafic and silicic dikes or sills. Such rocks are absent in the central and western parts of the field. It is likely that deep, hot, hydrothermal brines recharge the reservoir through fault H (**Figure 21(d)** and **21(e)**). The hot fluids flow along a Tertiary geothermal reservoir comprising deltaic sandstones and shales and a thermal plume of hot brines close to boiling ascends through a gap in the shales, before flowing toward the west through this and an overlying Quaternary geothermal reservoir of similar depositional origin. The Quaternary geothermal reservoir is connected to shallow aquifers by fault L, and fluid discharges occur from these in an arc west and southwest of the field, forming hot springs and fumaroles at the surface and shallow zones of temperature reversals. To the east, a zone of cold water recharge overlies the hot plume.

A map of the SP anomalies measured over the field was presented by Fitterman and Corwin (1982), from a field survey conducted by Corwin et al. (1979); see also Corwin and Hoover (1979). This map covers an area of more than 300 km^2 (**Figures 21(c)** and **22(a)**). The interpolated data exhibit a dipolar anomaly of ±80 mV with a peak-to-trough distance of 7 km. The SP data were acquired in 1977–78, shortly after the beginning of geothermal power production in 1973. To explain the observed SP anomaly, Fitterman and Corwin (1982) assumed a thermoelectric source mechanism located within a fault separating regions of different thermoelectric properties. However, Jardani et al. (2008) invoked a predominantly streaming source mechanism, associated with convection of the groundwater. They interpreted several existing 2-D resistivity surveys to determine the large-scale resistivity structure of the Cerro Prieto geothermal field, before inverting the surface SP data, using an approach similar to those outlined in **Section 11.09.4.4**, to determine the 3-D current sources. The upwelling plume of hot groundwater rising to the ground surface in the central part of the investigated area can clearly be observed in the inverted current source data, as can the discharge to the ground surface in the southwest part (**Figure 22(b)**). Thus, inversion of surface SP data, in conjunction with resistivity data, can be used to image 3-D subsurface flow.

11.09.6 Future Directions

Classical applications of the SP method focus almost exclusively on mineral exploration and borehole logging. If this was to continue, then the SP method would die away. However, a whole host of new applications, in areas of environmental and engineering site surveying, hydrology, monitoring of volcanic and seismic activity, leakage detection from dams and embankments, and location and monitoring of contaminant plumes, have seen it reinvigorated. In part, these new applications follow from recognition that the SP method often responds directly to subsurface processes of interest; they also reflect new understanding of the underlying physical processes, developments in hardware that allow data to be acquired more reliably using permanent as well as roving survey electrodes, and developments of quantitative as well as qualitative interpretation methods. The SP method is not yet as widely used as, for example, resistivity surveying, but it is moving away from its (undeserved) reputation as a geophysical relic or oddity, of historical interest but little current utility. Future directions are likely to include the development of joint inversion algorithms for complementary geophysical data, and application to new problems such as monitoring of subsurface flow in hydrocarbon reservoirs and during CO_2 sequestration. In this context, the future looks bright.

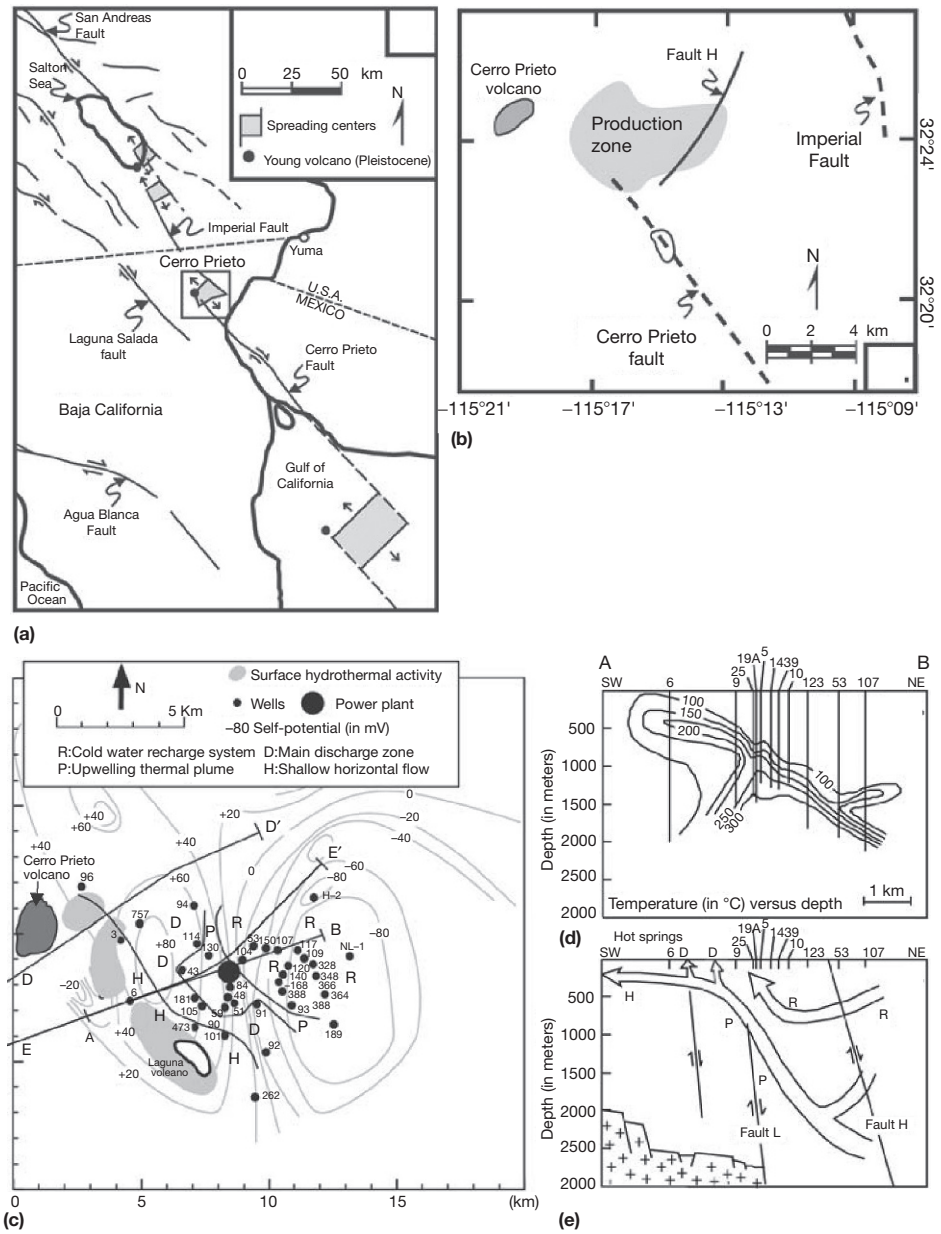


Figure 21 SP anomaly and subsurface temperature and flow patterns in the Cerro Prieto geothermal field. (a) Regional- and (b) small-scale maps of the field. (c) Map view showing the main flow regimes, and the SP contours acquired by Fitterman and Corwin (1982). Numbered dots denote boreholes. The profiles EE′ and DD′ are examples of DC resistivity (dipole–dipole) profiles used to determine the 3-D resistivity structure. (d) Cross section along profiles A–B in map (c), showing subsurface temperature distribution based on the calcite–water oxygen isotope geothermometer. The temperature distribution shows a broad area of temperature inversion in the southwest due to a zone of horizontal hot brine flow. (e) The same cross section showing main flow regimes. Modified from Jardani A, Revil A, Boléve A, and Dupont JP (2008) 3D inversion of self-potential data used to constrain the pattern of groundwater flow in geothermal fields. *Journal of Geophysical Research* 113: B09204. http://dx.doi.org/10.1029/2007JB005302.

11.09.7 Summary

The SP method comprises the passive measurement of electric potential at the ground surface and in boreholes. SP methods have a number of advantages over other geophysical techniques: They are often cheaper and quicker to implement, requiring only a pair (or more) of suitable electrodes and a high-impedance voltmeter, and data can be obtained over large regions with dense sampling in both space and time. Moreover, SP anomalies are often directly related to the process of interest, such as changes in groundwater flow, chemistry, and/or temperature. The disadvantages largely lie in interpreting the

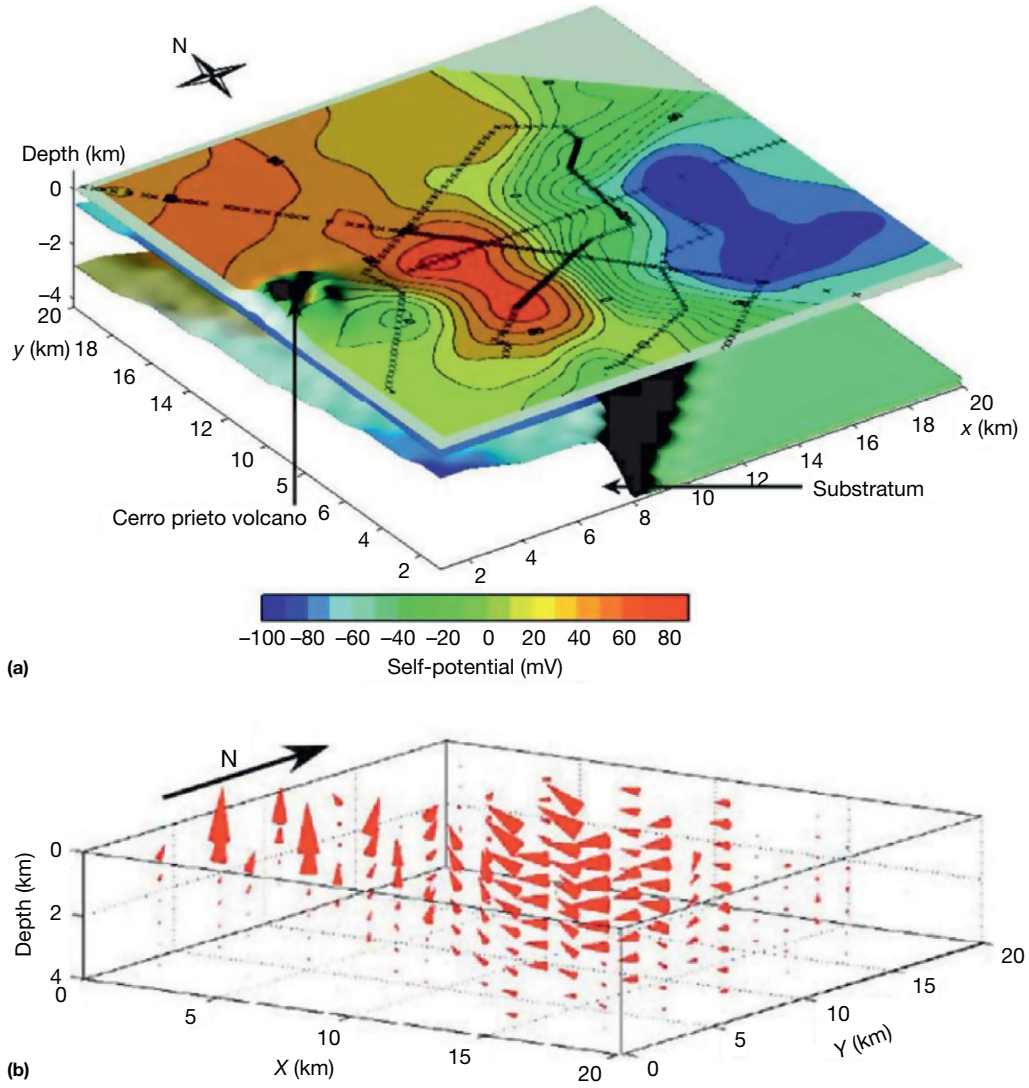


Figure 22 (a) SP data at the ground surface of the Cerro Prieto geothermal field (data from Fitterman and Corwin, 1982). The location of the SP measurements is indicated by the crosses. The position of the substratum (granite and metasedimentary upper Cretaceous basement) is determined from borehole data plus the seismic and deep resistivity surveys. (b) 3-D inversion of the direction of the streaming current density. The size of the arrows is proportional to the magnitude of the current density. The area of hydrothermal discharge in the southwest part of the investigated area (see areas D and H in **Figure 21(c)** and **21(e)**) can be seen. Modified from Jardani A, Revil A, Bolève A, and Dupont JP (2008) 3D inversion of self-potential data used to constrain the pattern of groundwater flow in geothermal fields. *Journal of Geophysical Research* 113: B09204. http://dx.doi.org/10.1029/2007JB005302.

data, which can be more challenging than other geophysical techniques. Similar to gravity and magnetic methods, SP measurements are purely passive, so there is often no way of adjusting source parameters to help identify signals of interest. Moreover, SP signals arise from a variety of sources, and distinguishing these can be challenging. Traditional SP surveys for mineral exploration, and borehole SP logs, have been interpreted qualitatively or semiquantitatively; however, a new generation of inversion techniques for SP measurements are now becoming available, driven by improved understanding of the underlying physical processes and increased computing power. Furthermore, the range and number of applications of the SP methods have rapidly increased in recent years.

References

Agar JN, Mou CY, and Lin JL (1989) Single-ion heat of transport in electrolyte solutions: A hydrodynamic theory. *Journal of Physical Chemistry* 93(5): 2079–2082.

Ahmed S, Jardani A, Revil A, and Dupont JP (2013) SP2DINV: A 2D forward and inverse code for streaming potential problems. *Computers and Geosciences* 59: 9–16.

Arora T, Linde N, Revil A, and Castermant J (2007) Non-intrusive characterization of the redox potential of landfill leachate plumes from self-potential data. *Journal of Contaminant Hydrology* 92: 274–292.

Atekwana E, Werkema DD, and Atekwana EA (2006) Biogeophysics: The effects of microbial processes on geophysical properties of the shallow subsurface. *Applied Hydrogeophysics* 71: 161–193.

Aubert M and Atangana QY (1996) Self-potential method in hydrogeological exploration of volcanic areas. *Ground Water* 34: 1010–1016.

Beattie JK and Djerdjev AM (2004) The pristine oil/water interface: Surfactant-free hydroxide-charged emulsions. *Angewandte Chemie* 116: 3652–3655. http://dx.doi.org/10.1002/ange.200453916.

Bigalke J and Grabner EW (1997) The geobattery model: A contribution to large-scale electrochemistry. *Electrochimica Acta* 42: 3443–3452.

Birch FS (1993) Testing Fournier's method for finding water table from self-potential. *Ground Water* 31: 50–56.

Birch FS (1998) Imaging the water table by filtering self-potential profiles. *Ground Water* 36: 779–782.

Bogoslovsky VA and Ogilvy AA (1973) Deformations of natural electric fields near drainage structures. *Geophysical Prospecting* 21: 716–723.

Bolève A, Revil A, Janod F, Mattiuzzo JL, and Fry J-J (2009) Preferential fluid flow pathways in embankment dams imaged by self-potential tomography. *Near Surface Geophysics* 7(5): 447–462. http://dx.doi.org/10.3997/1873-0604.2009012.

Braun BM and Weingartner H (1985) Transference numbers of aqueous NaCl and Na_2SO_4 at 25 °C from EMF measurements with sodium-selective glass electrodes. *Journal of Solution Chemistry* 14: 675–686.

Castermant J, Mendonça CA, Revil A, Trolard F, Bourrié G, and Linde N (2008) Redox potential distribution inferred from self-potential measurements associated with the corrosion of a burden metallic body. *Geophysical Prospecting* 56: 269–282. http://dx.doi.org/10.1111/j.1365-2478.2007.00675.x.

Corry CE (1985) Spontaneous polarization associated with porphyry sulfide mineralization. *Geophysics* 50: 1020–1034.

Corry CE, DeMoully GT, and Gerety MT (1983) *Field Procedure Manual for Self-Potential Surveys*. Australia: Zonge Engineering Research Organization.

Corwin RF (1989) Data quality for engineering self-potential surveys. In: *Detection of Subsurface Flow Phenomena*. *Lecture Notes in Earth Sciences*, vol. 27, pp. 49–72. Berlin: Springer.

Corwin RF (1990) The self-potential method for environmental and engineering applications. In: Ward SW (ed.) *Geotechnical and Environmental Geophysics*, vol. I. Society of Exploration Geophysicists.

Corwin RF and Hoover BH (1979) The self-potential method in geothermal exploration. *Geophysics* 44: 226–245.

Corwin RF, Morrison HF, Diaz S, and Rodriguez J (1979) Self potential studies at the Cerro Prieto geothermal field, 20–22 September 1978, pp. 204–210. *Report LBL-7098*. San Diego: Lawrence Berkeley Lab.

Darnet M, Maineult A, and Marquis G (2004) On the origin of self-potential (SP) anomalies induced by water injection into geothermal reservoirs. *Geophysical Research Letters* 31: L19609. http://dx.doi.org/10.1029/2004GL020922.

Darnet M, Marquis G, and Sailhac P (2003) Estimating aquifer hydraulic properties from the inversion of surface streaming potential (SP) anomalies. *Geophysical Research Letters* 30: 12–21.

Doherty R, Kulessa B, Ferguson AS, Larkin MJ, Kulakov LA, and Kalin RM (2010) A microbial fuel cell in contaminated ground delineated by electrical self-potential and normalized induced polarization data. *Journal of Geophysical Research: Biogeosciences (2005–2012)* 115.

Doll HG (1948) The SP Log: Theoretical Analysis and Principles of Interpretation. *Transactions of the American Institute of Mining, Metallurgical and Petroleum Engineers* 179: 146–186.

Dorfman MH, Oskay MM and Gaddis MP (1977) Self-potential profiling – A new technique for determination of heat movement in a thermal oil recovery flood. SPE Paper 6790, presented at the 1977 *Annual Conference and Exhibition of the SPE*.

Doussan C, Jouniaux L, and Thony J-L (2002) Temporal variations of SP and unsaturated water flow in loam and clay soils: A seasonal field study. *Journal of Hydrology* 267: 173–185.

Ernstson K and Scherer HU (1986) Self-potential variations with time and their relation to hydrogeologic and meteorological parameters. *Geophysics* 50: 1967–1977.

Fachin SJS, Abreu EL, Mendonça CA, Revil A, Novaes GC, and Vasconcelos SS (2012) Self-potential signals from an analog biogeobattery model. *Geophysics* 77(4): EN29–EN37. http://dx.doi.org/10.1190/geo2011-0352.1.

Finizola A, Lénat JF, Macedo O, Ramos D, Thouret JC, and Sortino F (2004) Fluid circulation and structural discontinuities inside Misti volcano (Peru) inferred from self potential measurements. *Journal of Volcanology and Geothermal Research* 135(4): 343–360.

Fitterman DV (1978) Electrokinetic and magnetic anomalies associated with dilatant regions in a layered Earth. *Journal of Geophysical Research* 83: 5923–5928.

Fitterman DV (1979) Theory of electrokinetic-magnetic anomalies in a faulted half-space. *Journal of Geophysical Research* 84: 6031–6041.

Fitterman DV (1981) Correction to 'Theory of electrokinetic-magnetic anomalies in a faulted half-space'. *Journal of Geophysical Research* 86: 9585–9588.

Fitterman DV and Corwin RF (1982) Inversion of self-potential data from the Cerro-Prieto geothermal field, Mexico. *Geophysics* 47: 938–945.

Fournier C (1989) Spontaneous potentials and resistivity surveys applied to hydrogeology in a volcanic area: Case history of the Chaîne des Puys (Puy-de-Dôme, France). *Geophysical Prospecting* 37: 647–668.

Fox RW (1830) On the electromagnetic properties of metalliferous veins in the mines of Cornwall. *Proceedings of the Royal Society of London* 2: 411.

Freund F, Salgueiro da Silva MA, Lau BWS, Takeuchi A, and Jones HH (2007) Electric currents along earthquake faults and the magnetization of pseudotachylite veins. *Tectonophysics* 431: 131–141.

Glover PWJ and Jackson MD (2010) Borehole electrokinetics. *The Leading Edge* 29: 724–728.

Glover PWJ, Walker E, and Jackson MD (2012) Streaming-potential coefficient of reservoir rock: A theoretical model. *Geophysics* 77: D17–D43, ISSN: 0016-8033.

Goldie M (2002) Self-potentials associated with the Yanacocha high-sulfidation gold deposit in Peru. *Geophysics* 67: 684–689. http://dx.doi.org/10.1190/1.1484511.

Gorby YA, Yanina S, McLean JS, Rosso KM, Moyles D, and Dohnalkova A (2006) Electrically conductive bacterial nanowires produced by Shewanella oneidensis strain MR-1 and other microorganisms. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11358–11363. http://dx.doi.org/10.1073/pnas.0604517103.

Guichet X, Jouniaux L, and Pozzi JP (2003) Streaming potential of a sand column in partial saturation conditions. *Journal of Geophysical Research Solid Earth* 108: 445–460. http://dx.doi.org/10.1029/2001JB001517.

Gulamali MZ, Leinov E, and Jackson MD (2011) Self-potential anomalies induced by water injection into hydrocarbon reservoirs. *Geophysics* 76: F283–F292.

Hearst JR, Nelson PH, and Paillett FL (2000) *Well Logging for Physical Properties*. Chichester: Wiley.

Herwanger J, Pain CC, Binley A, de Oliveira CRE, and Worthington MH (2004) Anisotropic resistivity tomography. *Geophysical Journal International* 158: 409–425.

Hubbard SS, Zhang JY, Monteiro PJM, Peterson JE, and Rubin Y (2003) Experimental detection of reinforcing bar corrosion using non-destructive geophysical techniques. *ACI Materials Journal* 100: 501–510.

Hunter RJ (1981) *Zeta Potential in Colloid Science*. New York: Academic Press.

Ikard SJ, Revil A, Jardani A, Woodruff WF, Parekh M, and Mooney M (2012) Saline pulse test monitoring with the self-potential method to nonintrusively determine the velocity of the pore water in leaking areas of Earth dams and embankments. *Water Resources Research* 48: W04201.

Ishido T, Kikuchi T, and Sugihara M (1989) Mapping thermally driven upflows by the self-potential method. In: Beck AE, et al. (ed.) *Hydrogeological Regimes and Their Subsurface Thermal Effects*. *Geophysical Monograph Series*, vol. 47 (IUGG vol. 2), pp. 151–158. Washington, DC: American Geophysical Union.

Ishido T and Mizutani H (1981) Experimental and theoretical basis of electrokinetic phenomena in rock-water systems and its applications to geophysics. *Journal of Geophysical Research* 86: 1763–1775.

Ishido T and Pritchett JW (1999) Numerical simulation of electrokinetic potentials associated with subsurface fluid flow. *Journal of Geophysical Research* 104: 15247–15259.

Jackson MD (2010) Multiphase electrokinetic coupling: Insights into the impact of fluid and charge distribution at the pore-scale from a bundle of capillary tubes model. *Journal of Geophysical Research* 115: B07206. http://dx.doi.org/10.1029/2009JB007092.

Jackson MD, Butler AP, and Vinogradov J (2012a) Measurements of spontaneous potential in chalk with application to aquifer characterisation in the southern UK. *Quarterly Journal of Engineering Geology and Hydrogeology* 45: 457–471. http://dx.doi.org/10.1144/qjegh2011-021.

Jackson MD, Gulamali MY, Leinov E, Saunders JH, and Vinogradov J (2012b) Spontaneous potentials in hydrocarbon reservoirs during waterflooding: Application to waterfront monitoring. *Society of Petroleum Engineers Journal* 17: 53–69.

Jackson MD and Leinov E (2012) Pore-level charge distribution and its impact on streaming currents in porous media. *International Journal of Geophysics* 2012: 1–12. http://dx.doi.org/10.1155/2012/897807.

Jackson MD and Vinogradov J (2012) Impact of wettability on laboratory measurements of streaming potential in carbonates. *Colloids and Surfaces A – Physicochemical and Engineering Aspects* 393: 86–95 ISSN: 0927-7757.

Jardani A, Dupont JP, and Revil A (2006) Self-potential signals associated with preferential ground water flow pathways in sinkholes. *Journal of Geophysical Research* 111: B09204. http://dx.doi.org/10.1029/2005JB004231.

Jardani A, Revil A, Bolève A, and Dupont JP (2008) 3D inversion of self potential data used to constrain the pattern of groundwater flow in geothermal fields. *Journal of Geophysical Research* 113: B09204. http://dx.doi.org/10.1029/2007JB005302.

Jardani A, Revil A, Santos F, Fauchard C, and Dupont JP (2007) Detection of preferential infiltration pathways in sinkholes using joint inversion of self-potential and EM-34 conductivity data. *Geophysical Prospecting* 55: 1–11. http://dx.doi.org/10.1111/j.1365-2478.2007.00638.x.

Jardani A, Revil A, Barrash W, et al. (2009) Reconstruction of the water table from self potential data: A Bayesian approach. *Ground Water* 47(2): 213–227.

Jordan JR and Campbell FL (1986) In: *Well Logging II – Electric and Acoustic Logging*, *SPE Monograph Society of Petroleum Engineers*, vol. 10, 159 pp.

Jouniaux L and Pozzi JP (1995) Streaming potential and permeability of saturated sandstones under triaxial stress: Consequences for electrotelluric anomalies prior to earthquakes. *Journal of Geophysical Research* 100(B6): 10197–10209.

Kato S, Hashimoto K, and Watanabe K (2012) Microbial interspecies electron transfer via electric currents through conductive minerals. *Proceedings of the National Academy of Sciences of the United States of America* 109: 10042–10046. http://dx.doi.org/10.1073/pnas.1117592109.

Kosmulski M and Dahlsten P (2006) High ionic strength electrokinetics of clay minerals. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 291: 212–218.

Leinov E and Jackson MD (2014) Experimental measurements of the SP response to concentration and temperature gradients in sandstones with application to subsurface geophysical monitoring. *Journal of Geophysical Research: Solid Earth* 119. http://dx.doi.org/10.1002/2014JB011249.

Leinov E, Vinogradov J, and Jackson MD (2010) Salinity dependence of the thermoelectric coupling coefficient in brine-saturated sandstones. *Geophysical Research Letters* 37: L23308. http://dx.doi.org/10.1029/2010GL045379.

Lide DR (ed.) (2009) *CRC Handbook of Chemistry and Physics*. Boca Raton, FL: CRC Press/Taylor and Francis, 89th edn.

Linde N and Revil A (2007) Inverting residual self-potential data for redox potentials of contaminant plumes. *Geophysical Research Letters* 34: L14302.

Linde N, Revil A, Bolève A, et al. (2007) Estimation of the water table throughout a catchment using self-potential and piezometric data in a Bayesian framework. *Journal of Hydrology* 334: 88–98. http://dx.doi.org/10.1016/j.jhydrol.2006.09.027.

Lynch EJ (1962) In: *Formation Evaluation*. *Harper's Geoscience Series,* New York: Harper and Row, 212 pp.

Maineult A, Bernabé Y, and Ackerer P (2005) Detection of advected concentration and pH fronts from self-potential measurements. *Journal of Geophysical Research* 110: B11205.

Maineult A, Strobach E, and Renner J (2008) Self-potential signals induced by periodic pumping tests. *Journal of Geophysical Research* 113: B01203.

Marquis G, Darnet M, Sailhac P, Singh AK, and Gerard A (2002) Surface electric variations induced by hydraulic stimulation: An example from the Soultz HDR site. *Geophysical Research Letters* 29: 7.1–7.4. http://dx.doi.org/10.1029/2002GL015046.

Marshall DJ and Madden TR (1959) Induced polarization, a study of its cases. *Geophysics* 25: 790–816.

Martinez-Pagan P, Jardani A, Revil A, and Haas A (2010) Self-potential monitoring of a salt plume. *Geophysics* 75: WA17–WA25.

Meiser P (1962) A method of quantitative interpretation of self potential measurements. *Geophysical Prospecting* 10: 203–218.

Mendonça CA (2008) Forward and inverse self-potential modeling in mineral exploration. *Geophysics* 73: F33–F43.

Michel S and Zlotnicki J (1998) Self potential and magnetic surveys. Tools to investigate faulting, fluid circulation and eruptive dynamism of volcanoes: Example of La Fournaise (Réunion Island). *Journal of Geophysical Research* 103: 17845–17857.

Minsley BJ, Sogade J, and Morgan FD (2007a) Three-dimensional source inversion of self-potential data. *Journal of Geophysical Research* 112: B02202. http://dx.doi.org/10.1029/2006JB004262.

Minsley BJ, Sogade J, and Morgan FD (2007b) Three-dimensional self-potential inversion for subsurface DNAPL contaminant detection at the Savannah River Site, South Carolina. *Water Resources Research* 43: W04429.

Moore JR, Glaser SD, Morrison HF, and Hoversten GM (2004) The streaming potential of liquid carbon dioxide in Berea sandstone. *Geophysical Research Letters* 31(17): L17610. http://dx.doi.org/10.1029/2004GL020774.

Murty BVS and Haricharan P (1985) Nomogram for the complete interpretation of spontaneous polarization profiles and sheet like and cylindrical two-dimensional sources. *Geophysics* 50: 1127–1135.

Naudet V, Revil A, Bottero J-Y, and Béegassat P (2003) Relationship between self-potential (SP) signals and redox conditions in contaminated groundwater. *Geophysical Research Letters* 30(21): 2091. http://dx.doi.org/10.1029/2003GL018096.

Naudet V, Revil A, Rizzo E, Bottero J-Y, and Bégassat P (2004) Groundwater redox conditions and conductivity in a contaminant plume from geoelectrical investigations. *Hydrology and Earth System Sciences* 8(1): 8–22.

Ntarlagiannis D, Atekwana E, Hill E, and Gorby Y (2007) Microbial nanowires: Is the subsurface 'hardwired'? *Geophysical Research Letters* 34: L17305. http://dx.doi.org/10.1029/2007GL030426.

Nyquist JE and Corry CE (2002) Self-potential: The ugly duckling of environmental geophysics. *The Leading Edge* 21: 446–451.

Ogilvy AA, Ayed MA, and Bogoslovsky VA (1969) Geophysical studies of water leakages from reservoirs. *Geophysical Prospecting* 22: 3662.

Ortiz I Jr., Osoba JS, and von Gonten WD (1973) Relationship of the electrochemical potential of porous media with hydrocarbon saturation. *The Log Analyst* 14: 25–32.

Pain CC, Herwanger JV, Saunders JH, et al. (2003) Anisotropic resistivity inversion. *Inverse Problems* 19: 1081–1111.

Patella D (1997) Introduction to ground surface self-potential tomography. *Geophysical Prospecting* 45: 653–681.

Paul MK (1965) Direct interpretation of self-potential anomalies caused by inclined sheets of infinite horizontal extensions. *Geophysics* 30: 418–423.

Perrier F and Morat P (2000) Characterization of electrical daily variations induced by capillary flow in the non-saturated zone. *Pure and Applied Geophysics* 157: 785–810.

Perrier F and Pant SR (2005) Noise reduction in long-term self-potential monitoring with travelling electrode referencing. *Pure and Applied Geophysics* 162: 165–179.

Perrier F, Petiau G, Clerc G, et al. (1997) A one-year systematic study of electrodes for long period measurements of the electric field in geophysical environments. *Journal of Geomagnetism and Geoelectricity* 49: 1677–1696.

Petiau G (2000) Second generation of lead–lead chloride electrodes for geophysical applications. *Pure and Applied Geophysics* 157: 357–382.

Petiau G and Dupis A (1980) Noise, temperature coefficient, and long time stability of electrodes for telluric observations. *Geophysical Prospecting* 28: 792–804. http://dx.doi.org/10.1111/j.1365-2478.1980.tb01261.x.

Poldini E (1938) Geophysical exploration by spontaneous polarization methods. *Mining Magazine London* 59: 278–282, 347–352.

Pride SR and Morgan FD (1991) Electrokinetic dissipation induced by seismic waves. *Geophysics* 56(7): 914–925.

Rao AD and Babu RHV (1984) Quantitative interpretation of self potential anomalies due to two-dimensional sheet like bodies. *Geophysics* 48: 1659–1664.

Reppert PM and Morgan FD (2003) Temperature-dependent streaming potentials: 2. Laboratory. *Journal of Geophysical Research Solid Earth* 108: 2547. http://dx.doi.org/10.1029/2002JB001755.

Revil A (1999) Ionic diffusivity, electrical conductivity, membrane and thermoelectric potentials in colloids and granular porous media: A unified model. *Journal of Colloid and Interface Science* 212: 503–522.

Revil A, Cary L, Fan Q, Finizola A, and Trolard F (2005) Self-potential signals associated with preferential ground water flow pathways in a buried paleo-channel. *Geophysical Research Letters* 32: L07401. http://dx.doi.org/10.1029/2004GL022124.

Revil A, Hermite D, Voltz M, et al. (2002) Self-potential signals associated with variations of the hydraulic head during an infiltration experiment. *Geophysical Research Letters* 29: 1106. http://dx.doi.org/10.1029/2001GL014294.

Revil A, Karaoulis M, Srivastava S, and Byrdina S (2013) Thermoelectric self-potential and resistivity data localize the burning front of underground coal fires. *Geophysics* 78(5): B259–B273.

Revil A, Leroy P, and Titov K (2005) Characterization of transport properties of argillaceous sediments: Application to the Callovo-Oxfordian argillite. *Journal of Geophysical Research* 110: B06202. http://dx.doi.org/10.1029/2004JB003442.

Revil A and Linde N (2006) Chemico-electromechanical coupling in microporous media. *Journal of Colloid and Interface Science* 302: 682–694.

Revil AN, Linde A, Cerepi D, Jougnot S, Matthäi, and Finsterle S (2007) Electrokinetic coupling in unsaturated porous media. *Journal of Colloid and Interface Science* 313(1): 315–327. http://dx.doi.org/10.1016/j.jcis.2007.03.037.

Revil A, Mendonça CA, Atekwana E, Kulessa B, Hubbard SS, and Bolhen K (2010) Understanding biogeobatteries: Where geophysics meets microbiology. *Journal of Geophysical Research* 115: G00G02. http://dx.doi.org/10.1029/2009JG001065.

Revil A, Naudet V, and Meunier JD (2004) The hydroelectric problem of porous rocks: Inversion of the water table from self-potential data. *Geophysical Journal International* 159: 435–444.

Revil A, Naudet V, Nouzaret J, and Pessel M (2003) Principles of electrography applied to self-potential electrokinetic sources and hydrogeological applications. *Water Resources Research* 39: 1114. http://dx.doi.org/10.1029/2001WR000916.

Revil A, Pezard PA, and Glover PWJ (1999) Streaming potential in porous media. 1. Theory of the zeta-potential. *Journal of Geophysical Research* 104: 20021–20031.

Rittgers JB, Revil A, Karaoulis M, Mooney MA, Slater LD, and Atekwana EA (2013) Self-potential signals generated by the corrosion of buried metallic objects with application to contaminant plumes. *Geophysics* 78: EN65–EN82. http://dx.doi.org/10.1190/GEO2013-0033.1.

Rizzo E, Suski B, Revil A, Straface S, and Troisi S (2004) Self-potential signals associated with pumping-test experiments. *Journal of Geophysical Research* 109: B10203. http://dx.doi.org/10.1029/2004JB003049.

Sailhac P, Darnet M, and Marquis G (2003) Electrical streaming potential measured at the ground surface: Forward modeling and inversion issues for monitoring infiltration and characterizing the vadose zone. *Vadose Zone Journal* 3: 1200–1206.

Sailhac P and Marquis G (2001) Forward and inverse modelling of SP anomalies caused by subsurface flow. *Geophysical Research Letters* 28: 1851–1854.

Sato M and Mooney HM (1960) The electrochemical mechanism of sulfide self-potentials. *Geophysics* 25: 226–249.

Saunders JH, Jackson MD, and Pain CC (2006) A new numerical model of electrokinetic potential response during hydrocarbon recovery. *Geophysical Research Letters* 33: L15316. http://dx.doi.org/10.1029/2006GL026835.

Saunders JH, Jackson MD, and Pain CC (2008) Fluid flow monitoring in oil fields using downhole measurements of electrokinetic potential. *Geophysics* 73: 165–180.

Saunders JH, Jackson MD, Pain CC, and Vinogradov J (2012) Streaming potentials in hydrocarbon reservoir conditions. *Geophysics* 77: E77–E90. http://dx.doi.org/10.1190/geo2011-0068.1.

Schlumberger C, Schlumberger M, and Leonardsen EG (1934) A new contribution to subsurface studies by means of electrical measurements in drill holes. *Transactions of AIME* 110: 273–282.

Segesman F (1962) New SP correction charts. *Geophysics* 27: 1.

Sen PN (1989) Influence of microgeometry on membrane potential of shaly sands. *Geophysics* 54(12): 1543–1553.

Sheffer MR and Oldenburg DW (2007) Three-dimensional modelling of streaming potential. *Geophysical Journal International* 169(3): 839–848.

Sill WR (1983) Self-potential modeling from primary flows. *Geophysics* 48: 76–86.

Smits LJM (1968) SP log interpretation in shaly sands. *Society of Petroleum Engineers Journal* 8(2): 123–136.

Stoll J, Bigalke J, and Grabner EW (1995) Electrochemical modelling of self-potential anomalies. *Surveys in Geophysics* 16: 107–120.

Strand S, Høgnesen EJ, and Austad T (2006) Wettability alteration of carbonates – Effects of potential determining ions (Ca and SO_4) and temperature. *Colloids and Surfaces A: Physiochemical and Engineering Aspects* 275: 1–10.

Telford WM, Geldart LP, and Sheriff RE (1990) *Applied Geophysics*, 2nd edn. Cambridge: Cambridge University Press.

Thomas EC (1976) The determination of Q_v from membrane potential measurements in shaly sands. *Journal of Petroleum Technology* 28(9): 1087–1096.

Thompson AH and Gist GA (1993) Geophysical applications of electrokinetic conversion. *The Leading Edge* 12: 1169–1173.

Timm F and Möller P (2001) The relation between electric and redox potential: Evidence from laboratory and field measurements. *Journal of Geochemical Research* 72: 115–128.

Titov K, Loukhmanov V, and Potapov A (2000) Monitoring of water seepage from a reservoir using resistivity and self-polarization methods: Case history of the Petergoph fountain water supply system. *First Break* 18: 431–435.

Trique M, Perrier F, Froidefond T, Avouac JP, and Hautot S (2002) Fluid flow near reservoir lakes inferred from the spatial and temporal analysis of the electric potential. *Journal of Geophysical Research* 107: 5.1–5.28. http://dx.doi.org/10.1029/2001JB000482.

Vilomet J-D, Angeletti B, Moustier S, et al. (2001) Application of strontium isotopes for tracing landfill leachate plumes in groundwater. *Environmental Science and Technology* 35: 4675–4679.

Vinogradov J, Jaafar MZ, and Jackson MD (2010) Measurement of streaming potential coupling coefficient in sandstones saturated with natural and artificial brines at high salinity. *Journal of Geophysical Research* 115: B12204. http://dx.doi.org/10.1029/2010JB007593.

Vinogradov J and Jackson MD (2011) Multiphase streaming potential in sandstones saturated with gas/brine and oil/brine during drainage and imbibition. *Geophysical Research Letters* 38: L01301.

Virkutyte J, Sillanpää M, and Latostenmaa P (2002) Electrokinetic soil remediation – Critical overview. *Science of the Total Environment* 289: 97–121.

Wallace WE (1968) Observations on the SP curve in general and offshore problems in particular. Presented at the *SPWLA Ninth Annual Logging Symposium*, June 23–26.

Westermann-Clark GB and Christoforou CC (1986) The exclusion-diffusion potential in charged porous membranes. *Journal of Electroanalytical Chemistry* 198: 213–231.

Woodruff WF, Revil A, Jardani A, Nummedal D, and Cumella S (2010) Stochastic Bayesian inversion of borehole self-potential measurements. *Geophysical Journal International* 183: 748–764.

Worthington AE and Meldau RF (1958) Departure curves for the self-potential log. *Journal of Petroleum Technology* 10: 1.

Wyllie MRJ (1951) An investigation of the electrokinetic component of the SP curve. *Transactions of AIME* 192: 1–11.

Zhang P and Austad T (2006) Wettability and oil recovery from carbonates: Effects of temperature and potential determining ions. *Colloids and Surfaces A: Physiochemical and Engineering Aspects* 279: 179–187.

This page intentionally left blank

11.10 Tools and Techniques: Active-Source Electromagnetic Methods

DV Fitterman, U.S. Geological Survey, Long Island City, NY, USA

© 2015 Elsevier B.V. All rights reserved.

11.10.1	**Introduction**	296
11.10.1.1	Subjects Covered	296
11.10.1.2	Characterization of Active-Source EM Methods	297
11.10.1.3	Advantages of EM over Other Near-Surface Methods	297
11.10.2	**Physical Properties of Geologic Materials**	298
11.10.2.1	Electrical Resistivity	298
11.10.2.1.1	Operational definition of resistivity	298
11.10.2.1.2	Role of pore fluids	298
11.10.2.1.3	Formation factor	299
11.10.2.1.4	Effect of clay minerals	299
11.10.2.1.5	Range of resistivities in near-surface environment	300
11.10.2.2	Dielectric Permittivity	301
11.10.2.3	Magnetic Permeability	301
11.10.3	**Governing Equations**	302
11.10.3.1	Maxwell's Equations	302
11.10.3.2	Constitutive Relationships	302
11.10.3.3	Boundary Conditions	302
11.10.3.4	Quasi-static Approximation	302
11.10.4	**Controlled-Source EM Methods**	303
11.10.4.1	Frequency-Domain Analysis	303
11.10.4.1.1	Magnetic dipole in free space	303
11.10.4.1.2	Mutual impedance	304
11.10.4.1.3	Frequency-domain homogenous half-space response	305
11.10.4.1.4	Induction number and skin depth	305
11.10.4.1.5	Small-induction number behavior of homogeneous half-space	306
11.10.4.1.6	Large-induction number behavior of homogeneous half-space	306
11.10.4.1.7	Frequency domain-layered half-space response	306
11.10.4.1.8	Small-induction number behavior of layered half-space	307
11.10.4.1.9	Large-induction number behavior of layered half-space	309
11.10.4.1.10	Frequency-domain response of nonlayered geometries	309
11.10.4.2	Time-Domain Analysis	309
11.10.4.2.1	Time-domain homogeneous half-space response	310
11.10.4.2.2	Dimensionless time-domain parameter	311
11.10.4.2.3	Early-stage and late-stage behavior of homogeneous half-space	312
11.10.4.2.4	Time-domain-layered half-space response	312
11.10.4.2.5	Theoretical basis for asymptotic behavior of layered half-space	313
11.10.4.2.6	Examples of layered half-space response	313
11.10.4.2.7	Time-domain response of nonlayered geometries	314
11.10.5	**Apparent Resistivity**	315
11.10.5.1	General Concept	315
11.10.5.2	Frequency-Domain Apparent Conductivity Function	315
11.10.5.3	Time-Domain Apparent Resistivity Function	317
11.10.6	**Interpretation**	319
11.10.6.1	Inversion	319
11.10.6.2	Noise in Electromagnetic Data	320
11.10.7	**Practical Aspects of Survey Design**	321
11.10.7.1	Sounding Versus Profiling	323
11.10.7.2	Array Types	324
11.10.7.3	Noncircular Transmitter Loops	324
11.10.8	**Applications and Examples**	324
11.10.8.1	Saltwater Intrusion Mapping	324
11.10.8.2	Coastal Bathymetry (AEM)	328

11.10.8.3 Sea-Ice Thickness Mapping 328
11.10.8.4 Clay Mapping 329
11.10.8.5 Airborne Mapping of UXO 330
11.10.9 Summary 331
References 331

11.10.1 Introduction

The basis of near-surface electromagnetic (EM) methods dates back to the unification of electromagnetic theory by James Clerk Maxwell in the late 1800s (Maxwell, 1873). The application of this theory to geophysical exploration lay dormant for over half a century until Sundberg (1931) described an EM method to map faults. Soon afterward, Sundberg and Hedström (1934) used the technique to explore for mineral deposits. It took 60 years more before near-surface EM methods started to come into vogue.

Sundberg's method was effective and popular in Europe and South Africa for mineral exploration; however, it had a drawback: The equipment was heavy and not very portable. It used a large horizontal loop or grounded-wire source powered by a 500 Hz audio oscillator. The receiver was an air coil attached to a sensitive voltmeter capable of determining signal strength and phase. The configuration and size of the equipment were not feasible for reconnaissance surveying, but useful for suspected targets based on geologic mapping. Smaller and more portable equipment, such as the Slingram equipment, made electromagnetic exploration easier to use and expanded the application of the method (Frischknecht et al., 1991; Keller and Frischknecht, 1966; Parasnis, 1979). A typical Slingram system consists of a transmitter and receiver coil connected by a reference cable to determine the phase of the received signal with respect to the transmitter current. Further advances in frequency-domain EM methods included measuring multiple frequencies to obtain information over a range of depths. Measuring the ratios of various field components eliminated the need for a reference between the transmitter and receiver, which was cumbersome when using a large transmitter loop (Ryu et al., 1972).

As equipment became smaller, more reliable, and available off-the-shelf, geophysicists started to think beyond mineral exploration to groundwater and environmental investigations (McNeill, 1990). These new targets, typically shallower than mineral deposits, were within tens to a few hundreds of meters of the surface and required better resolution than earlier mineral applications where a simple *bump detector* was often all that was needed.

The path described here, which begins with mineral exploration and leads to near-surface investigations, has focused on EM systems that use a continuous, usually sinusoidal, transmitter signal. A somewhat similar history can be told for systems that employ a transient transmitter signal. The development of transient EM systems started in Russia in the 1940s where targets were petroleum and mineral resources (Kaufman and Keller, 1983). The technique was introduced in the West in the 1970s where it was used for geothermal exploration (Jackson and Keller, 1972; Keller, 1971). Development during the 1970s and 1980s in Australia and Canada resulted in commercially available transient electromagnetic (TEM) equipment used initially for mineral exploration. It was a matter of time before the transient systems were used for groundwater and environmental targets (Abu Rajab and El-Naqa; 2013; Christiansen and Christensen, 2003; Dickinson et al., 2010; Ellefsen et al., 1998; Fitterman, 1986, 1987; Fitterman et al., 1991, 1999; Hatch et al., 2010; Powell et al., 2003; Sørensen and Søndergaard, 1999; Steuer at al, 2008; Yang et al., 1999), thereby completing a cycle similar to that of frequency-domain systems.

11.10.1.1 Subjects Covered

It may be helpful for readers to have a brief overview of the materials found in this chapter. The subject is active-source EM methods, that is, those that generate a primary EM field by means of a controlled current or voltage source. This is in contrast to passive EM systems that make use of naturally occurring EM fields. The goal of EM and electrical geophysical surveys is to determine the subsurface electrical resistivity structure and interpret it in terms of a geologic and/or hydrologic model. The resulting model can then be used to achieve a specific exploration objective, such as locating a geologic boundary, mapping a groundwater or mineral resource, or evaluating the source and extent of environmental contamination. To achieve this goal, the practitioner must be knowledgeable in several areas. First is the electrical nature of geologic materials. Herein is found a discussion of the range of electrical resistivities commonly encountered in geologic materials, which is very large, and the factors that control and influence resistivity. Second, we look at the mathematics governing EM fields, in general, and more specifically the more tractable quasi-static regime common to low-frequency EM methods. Working in the quasi-static range eliminates ground-penetrating radar (GPR) and other very-high-frequency methods from the discussion. The analysis will look at frequency-domain and time-domain systems for a homogeneous and layered half-space. The half-space analysis provides insights into the factors controlling the EM fields. This analysis is extended to 1-D-layered models, which are of great practical value. Some geologic situations, however, are not describable by 1-D models, and a brief introduction to nonlayered geometries is given.

The chapter then moves from theory to practical aspects starting with the design of an EM survey, stressing the importance of model studies to assess the likelihood of a particular method achieving success. The relative advantages and disadvantages of various measurement geometries for different targets are discussed. Having completed a survey, the data need to be put into a form that allows easy comparison from one location to another. This leads to the idea of apparent resistivity. While only a data transformation, apparent resistivity (or conductivity) is a starting point for developing a subsurface electrical model. Model development is accomplished these days by means of data inversion. A brief introduction to

inversion, a formalism by which model parameters are adjusted so that the EM response of the model agrees with the measured response, is given. The chapter concludes with examples of the application of ground and airborne EM methods to several groundwater, geotechnical, and environmental problems.

It is also worthwhile to briefly mention what will not be covered in this chapter in addition to GPR and other very-high-frequency EM methods. Methods that use as sources naturally occurring EM fields produced by lightning and magnetic storms will not be discussed. These include magnetotellurics (MT) and audiomagnetotellurics (AMT). A related method, controlled-source AMT (CSAMT) is also not discussed. This method typically measures orthogonal electric and magnetic fields as a function of frequency. The ratio of these fields forms an impedance that can be transformed into an apparent resistivity. Lower frequencies give information to greater depths. These are well-developed techniques and should be considered for deeper targets in the range of 1–10 km, which are often beyond the depth of exploration of some active-source methods. Active-source, ocean-bottom EM methods have also been developed employing electric dipole sources (Chave et al., 1991). These methods, which are often referred to as controlled-source EM (CSEM), are a very active research topic. During the past decade, CSEM has been used extensively for deepwater oil exploration (Constable, 2006; Constable & Srnka, 2007). These techniques will not be discussed here.

11.10.1.2 Characterization of Active-Source EM Methods

The number of EM methods is large with many variations on a theme. All methods, however, can be classified by a few characteristics starting with the source of the electromagnetic energy that is applied to the target. This chapter is confined to active-source methods. Such methods have an electric current source or transmitter. The source generates a primary EM field either by passing current through a loop of wire or by injecting current into the ground through a pair of electrodes. Loop sources are called *inductive* sources, while electrode injection sources are called *galvanic* sources. Only inductive sources will be discussed. A further restriction on the scope of this chapter is that only low-frequency EM methods will be considered; that is, methods such as GPR and techniques described by EM wave propagation are outside our scope of interest.

The time-dependent nature of the source current can be used to further classify an EM method. Frequency-domain electromagnetic (FEM) methods use a sinusoidally varying current. Sources that simultaneously transmit a small number of sine waves are members of this class. Methods that use a transient source, such as a step or impulse function, are called time-domain or TEM methods. In practice, time-domain current waveforms are composed of multiple turn-on and turnoff steps sequenced to produce an alternating polarity square wave with a duty cycle of 50%.

EM methods can also be characterized by the quantity that is measured. As with the source, the sensor can be either inductively or galvanically coupled to the ground. Only inductively coupled sensors will be considered here. Generally, the sensors are small, multiturn coils attached to a voltmeter. Induction coil sensors measure a signal proportional to the time derivative of the magnetic field. Magnetic field sensors, such as SQUID magnetometers, have been less commonly used for EM systems. Frequency-domain systems typically measure the amplitude and phase of the signal or equivalently the *inphase* and *quadrature* (90° phase-shifted) components of the signal. Time-domain systems measure the signed magnitude of the signal. In this chapter, we will consider FEM and TEM methods that use inductively coupled sources and sensors.

EM systems were first developed for use on the ground. The inaccessibility and large scale of some exploration targets led to the development of airborne geophysical methods. Following the earlier success of aeromagnetic surveying (Hanna, 1990; Muffly, 1946), airborne EM systems started to appear in the early 1950s (Barringer, 1990; Becker et al., 1990). By the early 1970s, helicopter-borne EM systems were developed that carried the transmitter and receiver in an instrumentation pod, called a *bird*, slung below the helicopter (Fraser, 1972, 1979). Examples of both ground-based and airborne EM methods will be presented. Ground-based EM systems have been floated over bodies of water to investigate below the water body (Fitterman, 2014; Fitterman et al., 1999; Goldman et al., 1996).

11.10.1.3 Advantages of EM over Other Near-Surface Methods

EM methods have some characteristics including the method of coupling to the ground, the physical response to pore fluids, and the speed of data acquisition that is not shared with other near-surface geophysical methods. These characteristics give EM methods an advantage.

In dry areas, such as deserts, injecting current into the ground through an electrode, as is done with DC resistivity methods, can be difficult, requiring special electrodes and watering of the area around the electrodes to reduce contact resistance (Keller and Frischknecht, 1966). Contact resistance can also be a problem in frozen terrain. EM methods do not succumb to this problem when the source and sensor are inductively coupled to the ground.

EM and electrical (galvanic source) methods share a characteristic in that their response depends directly on the electrical properties of the pore fluid contained in subsurface geologic materials. As discussed later in the text and in the companion **Chapter 11.04**, the bulk resistivity of geologic materials depends upon the porosity, degree of saturation, pore fluid resistivity, and clay content. This sensitivity allows EM and electrical measurements to directly estimate water quality. Monitoring saltwater intrusion is a prime use of this capability. Changes in pore fluid salinity also cause changes in pore fluid density. Seismic and gravity methods are highly influenced by the density of pore material (water, oil, and gas); however, they are insensitive to small changes in salinity that are important for environmental issues. Knight and Endres (2005) reported that the seismic velocity of distilled water and seawater differs by only 3.4%. When these fluids are in a rock matrix, the effect on bulk seismic velocity is even smaller. By contrast, the resistivity ratio of freshwater to seawater is >500 (Hem, 1970), and the ratio of resistivities of rocks saturated with these fluids would be nearly the same. Magnetic geophysical methods are not influenced by the physical properties of commonly encountered pore fluids.

The last advantage of EM methods is their relative speed of data acquisition compared to DC resistivity and seismic methods. The compact size of some EM systems makes them very portable allowing near-surface EM surveys to be conducted by one or two people. DC resistivity methods require the installation of an array of electrodes. The array can have as few as four electrodes: two for the transmitter and two for the receiver. The geometry of the array is expanded to change the depth of investigation (Grant and West, 1965; Keller and Frischknecht, 1966; Zonge et al., 2005). Expansion of the array can take a lot of time and requires a lot of physical space. Some modern DC resistivity systems use arrays of tens of electrodes that are electrically switched to change the depth of investigation. Repeated installation of a large number of electrodes as the array is moved along a profile is very labor-intensive, and the physical space requirement is not insignificant. Seismic methods also involve labor-intensive field operations because of the large numbers of geophones and associated recording systems. Provided that the EM technique is appropriate for an investigation, it can offer a significant economic advantage over other geophysical techniques.

EM methods are not without their shortcomings. Foremost is their inability to accurately determine the resistivity of high-resistivity materials. This happens because at high resistivities, the EM response diminishes approaching that of free space. Electromagnetic measurements are also often more susceptible than other methods of electromagnetic noise, such as that produced by power lines.

11.10.2 Physical Properties of Geologic Materials

Three physical properties influence the electromagnetic response: electrical resistivity (ρ), dielectric permittivity (ε), and magnetic permeability (μ). Because the behavior of these properties in geologic materials is different from other settings, a short discussion of them follows. More detail on these physical properties can be found in Chapter 11.04. Of these three properties, the electrical resistivity is the most diagnostic in determining the subsurface geology because of the wide range of values that are encountered in the near-surface environment. The resistivity of geologic materials can vary from about 0.1 Ω m to over 10 000 Ω m – a range of 10^5. Dielectric permittivity spans a smaller range, varying by a factor of about 8. The range of magnetic permeability is smaller yet varying by a factor of 5.

11.10.2.1 Electrical Resistivity

One of the goals of EM surveys is to collect data that can be converted into models of subsurface resistivity with the intent of interpreting these models in terms of geology, hydrologic conditions, environmental contamination, mineral resources, or other characteristics of interest. In order to achieve this goal, it is necessary to understand the factors controlling the electrical resistivity of geologic materials. The literature on this subject is extensive, ranging from laboratory and field observations to empirical relationships and theoretical, effective-media theories. The intent of this section is not to provide a comprehensive review of the topic, or even a detailed introduction, but rather to summarize the main factors controlling the electrical resistivity of geologic materials. The companion Chapter 11.04 provides a detailed introduction to the topic. Other sources of interest include Keller and Frischknecht (1966), Hearst and Nelson (1985), Olhoeft (1985), Keller (1988), Palacky (1988), and Knight and Endres (2005).

11.10.2.1.1 Operational definition of resistivity

An operational definition of resistivity for a solid is a useful starting point. Consider a right cylinder of length L and constant cross-sectional area A composed of a homogeneous, electrically conductive material. The electrical resistance between the flat ends of the cylinder will be proportional to L and inversely proportional to A:

$$R = \rho \frac{L}{A} \quad [1]$$

where the resistivity ρ is the constant of proportionality. Solving eqn [1] for the resistivity gives

$$\rho = R \frac{A}{L} \quad [2]$$

The unit of resistivity is ohm-meter (Ω m). Often, the reciprocal unit of electrical conductivity σ siemens per meter (S m^{-1}) is used where

$$\sigma = 1/\rho \quad [3]$$

In the hydrologic literature, water conductivity is usually expressed as specific conductance (SC) in units of microsiemens per centimeter (μS cm^{-1}), which is related to conductivity and resistivity through

$$\mathrm{SC} = 10^4 \sigma = \frac{10^4}{\rho} \quad [4]$$

11.10.2.1.2 Role of pore fluids

Common rock-forming silicate minerals are insulators at near-surface temperatures with resistivities in the range of 10^{12}–10^{17} Ω m (Grant and West, 1965; Keller and Frischknecht, 1966; Knight and Endres, 2005). Geophysically significant electric conduction can only take place by filling the pore space with conductive material such as water, thereby increasing conductivity by many orders of magnitude. Much of the conductivity of groundwater results from the dissolution of ionic solids giving rise to ions, the most common of which are Na^+, Mg^{2+}, Ca^{2+}, Cl^-, HCO^{3-}, and SO_4^{2-} (Freeze and Cherry, 1979; Hem, 1970). In general, the conductivity of pore water is linearly proportional to the dissolved ion concentration (Hem, 1970). The effective conductivity of a pore can be greater than that due to the pore water alone. Because of unsatisfied pH-dependent charges on the surface of silicate mineral grains, an excess of counterions accumulates in the pore space at the interface between the mineral and the fluid to provide overall electric neutrality of the rock. Under most natural geochemical conditions, the net surface charge is negative and the net counterion charge is positive. This excess charge gives rise to an enhanced conductivity called *surface conductivity*. To a first approximation, the surface conductivity forms a conduction network in parallel with the pore space. The contribution of surface conductivity to overall

conductivity is increased for small pore sizes and low pore fluid conductivity (Keller and Frischknecht, 1966; Knight and Endres, 2005). Surface conduction can be important when pore spaces are small and pore water is very fresh. Its effect on rock resistivity is similar to the presence of clay in pore spaces, which is discussed in the succeeding text.

Data from the Biscayne aquifer of south Florida illustrate the relationship of water resistivity (or specific conductance) to chloride content. The Biscayne aquifer is the primary source of drinking water along the east coast of south Florida. The quality of water in the aquifer ranges from fresh inland away from the coast to saline near the shore. Chloride concentration in the nearby Biscayne Bay can often exceed the 19 400 mg l^{-1} value of ocean water (Stumm and Morgan, 1996; Turekian, 1976) because of evaporation and limited circulation in the bay. Chloride values of greater than 1000 mg l^{-1} in groundwater are attributed to seawater intrusion and indicate nonpotable water. **Figure 1** is a graph of the resistivity and specific conductance of pore water from the Biscayne aquifer as a function of chloride content (Fitterman et al., 2012). Specific conductance has a power law dependence on chloride content except at very low chloride levels.

11.10.2.1.3 Formation factor

The relationship between formation (bulk rock) resistivity, pore fluid resistivity, and porosity was established by Archie (1942), who measured the resistivity of cores from Gulf Coast oil wells saturated with brine. Archie determined that the resistivity of the brine-saturated rock ρ_F was related to the brine resistivity ρ_W through the equation

$$\rho_F = F\rho_W \quad [5]$$

where F is called the formation factor. Archie found that for brine concentrations of 20 000–100 000 mg l^{-1} NaCl and porosities in the range of 0.1–0.4, F is constant for any given formation. **Figure 2** shows formation resistivity and pore water resistivity from wells in the Biscayne aquifer (Fitterman et al., 2012). The formation resistivities were determined from induction logs in screened intervals where water samples were collected. A formation factor of 5.1 was determined by regression.

When the formation factor is plotted as a function of porosity for samples from the same formation, the data are fit with a power law relationship commonly referred to as Archie's equation:

$$F = \frac{\rho_F}{\rho_W} = \phi^{-m} \quad [6]$$

The exponent m is called the cementation factor and varies with rock type. For sedimentary rocks, m is in the range of 1.8–2.0; for unconsolidated sands with no clay, the exponent is lower (1.4–1.6) (Jackson et al., 1978). Archie's equation even holds for crystalline rocks with porosity less than 0.01 provided the pore fluid is very conductive so that surface conduction effects are not important (Brace and Orange, 1968). For crystalline rocks, m is close to 2.0. Some studies have suggested the addition of a constant factor on the right-hand side of eqn [6] as a way of improving the fit to observed data (Knight and Endres, 2005; Winsauer et al., 1952; Wyllie and Gregory, 1953). Such a constant has no physical significance.

There are several situations where Archie's equation needs to be modified. Partial saturation of the pore space through the introduction of gas, oil, or other nonconducting fluid is the first. Partial saturation effectively increases the resistivity of the pore fluid. This is handled in eqn [6] by the inclusion of a term S^{-n} on the right-hand side, where S is the saturation and n is a constant, typically 2 ± 0.5 (Hearst and Nelson, 1985) to give

$$F = \frac{\rho_F}{\rho_W} = \phi^{-m} S^{-n} \quad [7]$$

11.10.2.1.4 Effect of clay minerals

The presence of clay minerals in the pore space is the second situation that requires a modification of Archie's equation. Waxman and Smits (1968) developed a model based on the assumption that clay in the pore spaces can be viewed as a parallel conduction network that augments conduction through the pore fluid. The clay effectively increases the

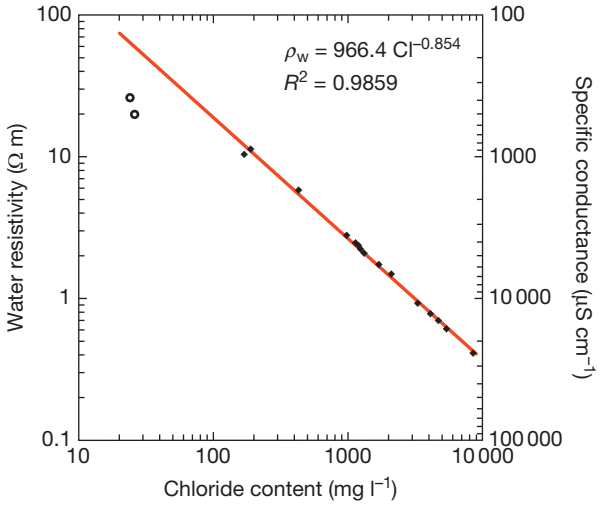


Figure 1 Electrical resistivity and specific conductance of water samples from the Biscayne aquifer of south Florida as a function of chloride content (Fitterman et al., 2012). The open circles are data points not used in the regression. These data probably fall below the regression line because of differences in water chemistry.

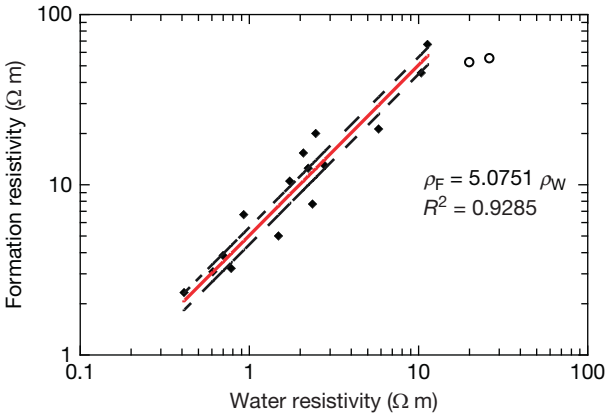


Figure 2 Formation resistivity as a function of water resistivity from wells in the Biscayne aquifer of south Florida (Fitterman et al., 2012). The equation describing the linear regression is shown. The dashed lines show the uncertainty in the formation resistivity estimate. The open circles are data points not used in the regression.

conductivity of the pore space. It is commonly assumed that the formation factors of the two conduction networks are the same, and the formation conductivity ($\sigma_F = 1/\rho_F$) is given by

$$\sigma_F = \frac{\sigma_W + \sigma_e}{F} = \frac{\sigma_W + BQ_v}{F} \quad [8]$$

where σ_e is the additional conductivity due to the exchangeable cations in the clay, B (S m^{-1} l Eq^{-1}) is an experimentally determined fitting parameter (Waxman and Thomas, 1974)

$$B = 3.83[1 - 0.83\exp(-0.5/\rho_W)] \quad [9]$$

and Q_v (Eq l^{-1}) is the effective valence concentration associated with the clay. Q_v is related to the cation exchange capacity (CEC) (mEq 0.01 g^{-1}) and the grain density of the rock ρ_{ma} (g l^{-1}) by

$$Q_v = \frac{(1-\phi)\rho_{ma}\mathrm{CEC}}{100\phi} \quad [10]$$

Q_v is determined using laboratory measurements of CEC and ρ_{ma}. For reference purposes, Hearst and Nelson (1985) described a *high clay content* sample as having a Q_v value of 1.530 Eq l^{-1}, while a *clay present* sample has a value of 0.239 Eq l^{-1}. Expressing eqn [8] as a resistivity gives

$$\rho_F = \frac{F}{1 + BQ_v\rho_W}\rho_W = F_E\rho_W \quad [11]$$

where

$$F_E = \frac{1}{1 + BQ_v\rho_W} \quad [12]$$

is the effective formation factor taking into account the presence of the clay.

Figure 3 shows plots of formation resistivity as a function of pore water resistivity for varying values of Q_v and a fixed formation factor of 5. Q_v equal to zero corresponds to no clay and plots as a straight line. Increasing clay content and Q_v causes the curves to depart from the $Q_v = 0$ line, with larger values of Q_v producing bigger differences. The effective formation factor as a function of Q_v for different values of ρ_W is shown in Figure 4. When the pore water is conductive (small ρ_W), the effect of clay content is not noticeable even for large values of Q_v. As ρ_W increases, the presence of clay influences the formation factor at lower clay concentrations.

The electrical resistivity of barren geologic materials, that is, those without metallic minerals, clays, carbonaceous materials, or organics, has very little frequency dependence (Olhoeft, 1985). Olhoeft showed a barren rock that exhibits less than 2% resistivity variation over a ten-decade range of frequency. On the other hand, the presence of metallic mineral grains gives rise to oxidation–reduction reactions, ion exchange processes take place when clays are present, and organic compounds will interact with clays when present. Any of these situations give rise to variations in resistivity with frequency. These variations can be diagnostic of the material causing the reaction, such as the tendency for oxidation–reduction reactions associated with the presence of metals, which produces the well-known induced polarization (IP) phenomenon used for mineral exploration and environmental studies (Angoran et al., 1974; Draskovits and Verö, 2005; Madden and Marshall, 1967; Marshall and Madden, 1959). Similarly, cation exchange processes can be used to map clays and clay–organic reactions to map hydrocarbon contaminants (Grimm et al., 2005; Olhoeft, 1985).

11.10.2.1.5 Range of resistivities in near-surface environment

The range of resistivities expected in the near-surface environment is depicted in Figure 5. The diagram is a modified form of one developed by Palacky (1988), whose focus was primarily mineral exploration. The purpose of the diagram is not to give

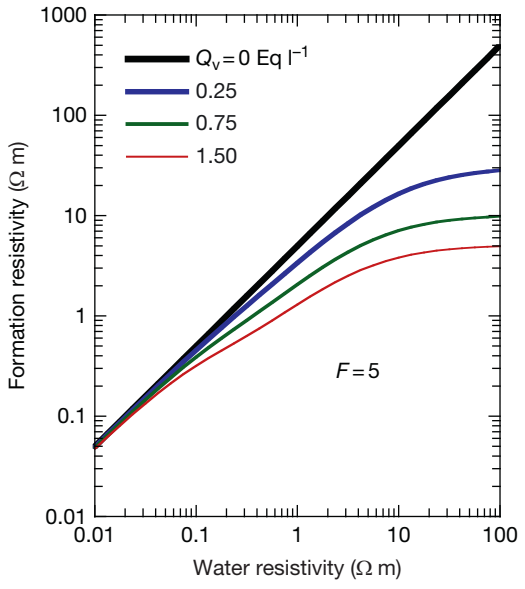


Figure 3 Formation resistivity based on the Waxman–Smits model as a function of water resistivity with varying values of effective clay valence concentration Q_v. Q_v equal to zero corresponds to no clay in the pore space. The formation factor is fixed at 5. As Q_v increases, the effective formation factor decreases.

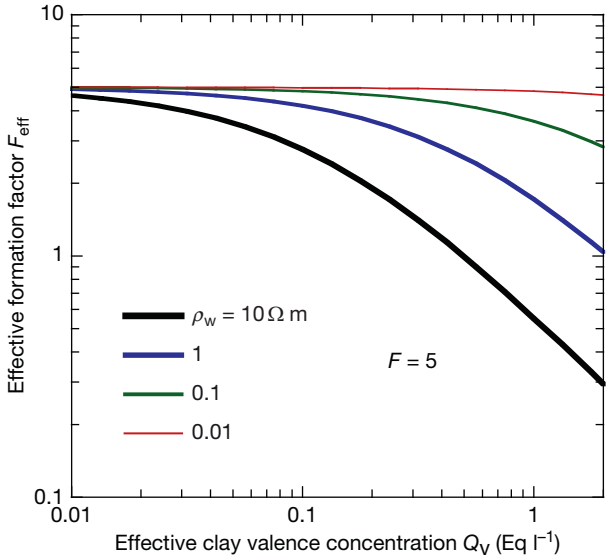


Figure 4 Effective formation factor as a function of effective clay valence concentration Q_v for different pore fluid resistivities. Increasing values of Q_v correspond to more clay in the pore space. The formation factor with no clay present is set to 5.

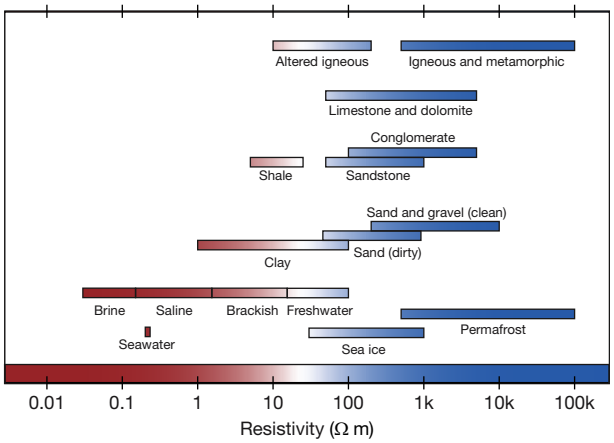


Figure 5 Range of resistivities expected in near-surface investigations. Diagram is based on data from Palacky (1988), Freeze and Cherry (1979), and Hearst and Nelson (1985).

specific resistivity values for geologic materials, but rather to give the range of values commonly encountered and to show the differences between rock types. The horizontal resistivity axis goes from very conductive (0.01 Ω m) to very resistive (100k Ω m), a span of seven orders of magnitude. Higher resistivities may be encountered; however, for electromagnetic measurements, values above 10k Ω m behave similarly.

The lower tier of resistivity ranges is for water, sea ice, and permafrost, of which water is the most important based on the nature of Archie's equation. Using an upper limit of 1000 mg l^{-1} total dissolved solids (TDSs) for drinking water, the minimum resistivity of freshwater is about 15 Ω m (Freeze and Cherry, 1979). The upper resistivity of freshwater has been put at 100 Ω m, which is several orders of magnitude less than that of distilled water. Also shown is the range of resistivities for brackish water, saline water, and brine, terms commonly used to categorize groundwater based on TDS levels (Freeze and Cherry, 1979). The other reference point is seawater with a resistivity of 0.2 Ω m (Turekian, 1976). When seawater freezes, the mobility of ions drastically decreases, and its resistivity increases as indicated by the sea-ice values. Permafrost also has high resistivity due to decreased ionic mobility. The conductivity of water has a temperature dependence of about +2% per °C controlled by the viscosity of the water. At temperatures above freezing, the resistivity of water-saturated rock increases as it is cooled. With further cooling, the resistivity discontinuously increases at the freezing point (Keller and Frischknecht, 1966). Permafrost has a similar temperature dependence as temperature nears the freezing point (Vanhala et al., 2009).

Unconsolidated materials compose the next tier of the diagram. Clay and sand and gravel are the end-members of this group. Clean sand, that is, sand without any clay, will have resistivities of more than 200 Ω m when freshwater-saturated. Decreasing water quality (and pore fluid resistivity) will decrease the formation resistivity values as described by Archie's equation. The resistivity of clay varies with the specific clay mineral present due to differences in CEC (Keller and Frischknecht, 1966). Clay resistivities in the range of 1–20 Ω m are common, however, in environments where heavy, annual rainfall has leached ions from the clay, resistivities can exceed 100 Ω m. The addition of clay to clean sand reduces the resistivity of the mixture, referred to here as *dirty sand*.

The third tier of resistivities is for sedimentary rocks. Two bands of resistivity are shown, one for shale and the other for sandstone and conglomerate, which correspond to the clay and sand and gravel in the unconsolidated tier. Resistivity values are controlled by porosity and clay content. Carbonate rocks (tier four) span a large range because of the wide variation in porosity that is possible.

The last tier is for unaltered and altered igneous and metamorphic rocks. The porosity of these rock types is usually very low, producing very high resistivities. Weathering and alteration lead to the formation of clay minerals in pore spaces, which lower the resistivity. The conductivity of the altered zone and type of clay minerals formed depend upon the parent rock type (Palacky, 1988). The less quartz contained in a rock, the more susceptible it is to alteration and the formation of conductive clays. Mafic rocks tend to alter to more conductive material than felsic rocks.

11.10.2.2 Dielectric Permittivity

Dielectric permittivity ϵ influences the magnitude of the electric field created by electric charges. The magnitude is inversely proportional to the dielectric permittivity. In free space, the permittivity is constant: $\epsilon_0 = 10^{-9}/36\pi\,\mathrm{F\,m^{-1}}$. In insulators and polarizable material, such as geologic materials, the dielectric permittivity increases. The relation between the dielectric of a material and that of free space is given by

$$\epsilon = \epsilon_r \epsilon_0 \qquad [13]$$

where ϵ_r is the *relative permittivity* or *dielectric constant*. Rock-forming minerals have dielectric constants in the range of 4–9 with the values for clays being slightly greater (9.5–13.7) (Knight and Endres, 2005). Water, which is ubiquitous in the near-surface environment, has a dielectric constant of 80. Knight and Endres (2005) gave an extensive review of dielectric mixing laws and effective-media theory that can be used to estimate the appropriate value of dielectric constant for a geologic situation. Dielectric permittivity does exhibit a decrease with frequency (Keller and Frischknecht, 1966; Knight and Endres, 2005); however, for our purpose, we will consider it to be constant.

11.10.2.3 Magnetic Permeability

The *magnetic permeability* of geologic materials is influenced by the presence of magnetic minerals. The magnetic permeability can be related to the free-space value $\mu_0 = 4\pi\,10^{-7}\,\mathrm{H\,m^{-1}}$ in a manner similar to that used for the dielectric permittivity:

$$\mu = \mu_r \mu_0 \qquad [14]$$

where μ_r is the *relative permeability*. Keller and Frischknecht (1966) gave values of the relative magnetic permeability for some common rock-forming minerals; values typically vary from 1 by less than 0.1%. Magnetic minerals, such as magnetite on the other hand, can have values as high as 5. Increasing the magnetite content of a rock increases its relative permeability.

Little information is available on the frequency dependence of magnetic permeability of geologic materials (Keller, 1988).

11.10.3 Governing Equations

Electromagnetic geophysical methods are based on the principles of electromagnetic induction as described by Maxwell's equations. The interaction between currents flowing in conductors and the magnetic fields that these currents produce is central to the process.

11.10.3.1 Maxwell's Equations

Maxwell's equations are as follows (Stratton, 1941):

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \quad [15]$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad [16]$$

$$\nabla \cdot \mathbf{D} = q \quad [17]$$

and

$$\nabla \cdot \mathbf{B} = 0 \quad [18]$$

Equation [15] is Ampere's law that describes the *magnetic field intensity* $\mathbf{H}$ ($\mathrm{A\,m^{-1}}$) produced by *current density* $\mathbf{J}$ ($\mathrm{A\,m^{-2}}$) and time variations of the *electric displacement* $\mathbf{D}$ ($\mathrm{C\,m^{-2}}$). Equation [16] is Faraday's law that describes the production of the *electric field* $\mathbf{E}$ ($\mathrm{V\,m^{-1}}$) by the time-varying *magnetic induction* $\mathbf{B}$ ($\mathrm{Wb\,m^{-2}}$). When conductors are present, these electric fields induce current flow in the conductors pierced by the magnetic field. Any induced currents try to produce a secondary magnetic field that counteracts the change in the primary magnetic field that induced the currents. Lenz's law can be used to determine the direction of the induced currents. Equation [17] is Gauss's law stating that the electric displacement is produced by *electric charge* q ($\mathrm{C\,m^{-3}}$). Finally, eqn [18] states that the magnetic induction $\mathbf{B}$ is solenoidal in form, that is, the lines of magnetic induction are closed and there are no magnetic monopoles. Additionally, we require the conservation of charge given by

$$\nabla \cdot \mathbf{J} = -\frac{\partial q}{\partial t} \quad [19]$$

Because we are not considering galvanically coupled sources, the right-hand side of eqn [19] will be zero, and the current density is divergence-free.

11.10.3.2 Constitutive Relationships

Two additional equations, called the *constitutive relationships*, are needed to relate the field intensity vectors to the flux density vectors:

$$\mathbf{D} = \epsilon \mathbf{E} \quad [20]$$

and

$$\mathbf{B} = \mu \mathbf{H} \quad [21]$$

In free space or media where ϵ and μ are known, the EM fields can be written in terms of $\mathbf{E}$ and $\mathbf{H}$ alone. Finally, Ohm's law will be used to relate the electric field to the current density it produces in conductors:

$$\mathbf{J} = \sigma \mathbf{E} \quad [22]$$

11.10.3.3 Boundary Conditions

The solution of Maxwell's equations depends upon the usual boundary conditions of the field (Ward and Hohmann, 1988). At an interface separating medium 1 from medium 2 with a unit normal at the interface $\hat{\mathbf{n}}$, the boundary conditions are as follows: The normal components of current density and magnetic induction are continuous

$$(\mathbf{J}_1 - \mathbf{J}_2) \cdot \hat{\mathbf{n}} = 0 \quad [23]$$

$$(\mathbf{B}_1 - \mathbf{B}_2) \cdot \hat{\mathbf{n}} = 0 \quad [24]$$

The normal component of electric displacement is discontinuous by an amount equal to the *surface charge density* Σ_s on the interface

$$(\mathbf{D}_1 - \mathbf{D}_2) \cdot \hat{\mathbf{n}} = \Sigma_s \quad [25]$$

The tangential components of electric and magnetic field are continuous

$$(\mathbf{E}_1 - \mathbf{E}_2) \times \hat{\mathbf{n}} = 0 \quad [26]$$

$$(\mathbf{H}_1 - \mathbf{H}_2) \times \hat{\mathbf{n}} = 0 \quad [27]$$

The solution of the field equations also depends upon the geometry of the conductors and the position and nature of the sources. Many illustrative examples can be found in standard electromagnetic texts (Griffiths, 1999; Jackson, 1962; Stratton, 1941). The geometries of interest to most geophysical problems, however, are very different from those discussed in physics textbooks. Geophysical problems are usually posed in an infinite half-space that may be homogeneous or layered and may include an embedded three-dimensional body. For an introduction to these situations, the works by Grant and West (1965), Keller and Frischknecht (1966), Ward (1967), Wait (1982), and Ward and Hohmann (1988) are highly recommended.

11.10.3.4 Quasi-static Approximation

One assumption that is often made to simplify the solution of Maxwell's equations is that the displacement currents are small compared with conduction currents. This is called the *quasi-static approximation*. The displacement currents are given by the second term on the right-hand side of eqn [15]. Thus, the quasi-static approximation requires that

$$\mathbf{J} \gg \frac{\partial \mathbf{D}}{\partial t} \quad [28]$$

When this condition is met, the magnetic field varies slowly enough that radiation effects can be ignored. Stated another way, it implies that the magnetic field has the same time dependence everywhere. For a given excitation frequency f, the free-space wavelength is c/f, where c is the speed of light ($3 \times 10^8\ \mathrm{m\,s^{-1}}$). Typical controlled-source EM systems operate in the frequency range from 100 Hz to 200 kHz with corresponding free-space wavelengths of 3000 to 1.5 km,

respectively. Provided the separation between transmitter and receiver is smaller than the free-space wavelength, the quasi-static approximation is justified, and the Biot–Savart law is appropriate to describe the magnetic field produced by current flow.

A different criterion is used in conductors to establish the validity of the quasi-static approximation. Assume that the magnetic and electric fields have an $e^{i\omega t}$ time dependence, then eqn [15] becomes

$$\nabla \times \mathbf{H} = (\sigma + i\omega\epsilon)\mathbf{E} \qquad [29]$$

after eliminating the time-dependent term from both sides of the equation. The displacement current term will be small when $\sigma \gg \omega\epsilon$. Assuming the extreme case of a very low conductivity ($\sigma = 10^{-4}\,\mathrm{S\,m^{-1}}$) and a large dielectric constant ($\epsilon = 10\epsilon_0\,\mathrm{F\,m^{-1}}$), displacement currents can be ignored at frequencies of less than 180 kHz. This cutoff frequency covers most electromagnetic techniques with the exception of GPR.

11.10.4 Controlled-Source EM Methods

In this section, we examine the behavior of frequency-domain and time-domain EM systems over a homogeneous and layered half-space. Before doing this, it is useful to understand what EM systems actually measure.

The current in the transmitter loop and the induced current in the ground produce magnetic and electric fields. The sensor used to measure these fields is either a small coil or grounded dipole attached to a voltmeter. We are limiting our scope to the former. The time-varying magnetic flux through the sensor (or receiver) coil produces a voltage described by the integral form of Faraday's law:

$$\begin{aligned} V &= \oint_{\text{coil}} \mathbf{E}\cdot d\mathbf{l} = -\frac{d\Phi}{dt} \\ &= -\frac{d}{dt}\int_{\text{coil}} d\Phi = -\frac{d}{dt}\int_{\text{coil}} \mathbf{B}\cdot\hat{\mathbf{n}}\,dA = -\mu_0\frac{d}{dt}\int_{\text{coil}} \mathbf{H}\cdot\hat{\mathbf{n}}\,dA \end{aligned} \qquad [30]$$

where $\Phi = \mathbf{B}\cdot\hat{\mathbf{n}}A$ is the magnetic flux, A is the area of the coil, and $\hat{\mathbf{n}}$ is the unit normal to the plane of the coil. For small coils, the magnetic field is constant over the coil and the voltage is given by

$$V = -\frac{dB_n}{dt}S_{Rx} = -\mu_0\frac{dH_n}{dt}S_{Rx} \qquad [31]$$

where the subscript n denotes the normal component of the field and S_{Rx} is the effective area of the receiver coil. For multi-turn coils, the effective area is the product of the coil area and the number of turns. Frequency-domain systems will have an $e^{i\omega t}$ time dependence that changes the voltage expression to

$$V(\omega) = -i\omega B_n S_{Rx} = -i\omega\mu_0 H_n S_{Rx} \qquad [32]$$

The receiver will resolve the voltage into inphase and quadrature components. In free space, the voltage is phase-shifted 90° with respect to the transmitter current, and nearby conductors will introduce additional, small phase shifts in the secondary magnetic field. It is customary to subtract the primary signal from the inphase component to give a secondary inphase signal. All of the quadrature signal is secondary. Both the secondary inphase and quadrature components are normalized by the magnitude of the primary field. The result is often reported as parts per million of the primary field. Time-domain systems measure only a real voltage.

11.10.4.1 Frequency-Domain Analysis

Many geologic situations can be described by a layered half-space, making analysis of this geometry of practical use. Consider a layered half-space with layer conductivities $\sigma_i = 1/\rho_i$ and layer thicknesses h_i (see Figure 6). The transmitter coil is located at the origin of a right-hand coordinate system with the z-axis pointing upward. The transmitter current I_0 flows in a circular loop of radius r_0. For distances R that are large compared to the loop radius, the current loop behaves like a magnetic dipole. The analysis will start with the field of a magnetic dipole in free space. Results for other geometries are typically normalized to the free-space result.

11.10.4.1.1 Magnetic dipole in free space

The primary field due to the magnetic dipole can be computed from the magnetic potential $\mathbf{A}$ using

$$\mathbf{H} = \frac{\nabla \times \mathbf{A}}{\mu_0} \qquad [33]$$

where the magnetic potential of the dipole is given by

$$\mathbf{A}(\mathbf{R}) = \frac{\mu_0}{4\pi}\frac{\mathbf{m}_{Tx} \times \mathbf{R}}{R^3} \qquad [34]$$

$\mathbf{R}$ is the vector from the dipole to the observation point, R is its magnitude, and $\mathbf{m}_{Tx}$ is the magnetic moment of the source (Griffiths, 1999; Reitz et al., 1979). The magnitude of the

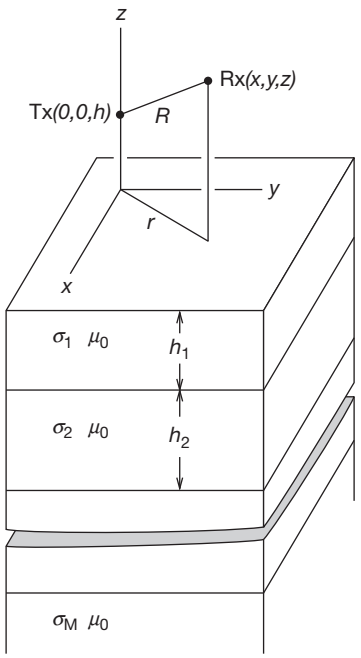


Figure 6 Layered half-space geometry and coordinate system. The location of centers of the transmitter (Tx) and receiver (Rx) loops is shown.

magnetic dipole is $m_{Tx}=S_{Tx}I_0=\pi r_0^2 I_0$. The magnetic dipole is oriented perpendicular to the plane of the circular current flow, and the magnitude of the current flow is I_0. Substituting eqn [34] into eqn [33] gives the magnetic field

$$\mathbf{H}=\frac{1}{4\pi}\left[\frac{3\mathbf{m}_{Tx}\cdot\mathbf{R}}{R^5}\mathbf{R}-\frac{\mathbf{m}_{Tx}}{R^3}\right] \quad [35]$$

Using the geometry shown in Figure 6, the transmitter loop is located at $(0,0,h)$, and the receiver loop is at (x,y,z). By placing the transmitter and receiver coils above the half-space, the analysis is applicable to airborne EM systems. The field components for the vertical magnetic dipole are given by

$$H_x^P=\frac{m_{Tx}}{4\pi}\frac{3x(z-h)}{R^5} \quad [36]$$

$$H_y^P=\frac{m_{Tx}}{4\pi}\frac{3y(z-h)}{R^5} \quad [37]$$

and

$$H_z^P=\frac{m_{Tx}}{4\pi}\left[\frac{3(z-h)^2}{R^5}-\frac{1}{R^3}\right] \quad [38]$$

where $R^2=r^2+(z-h)^2$ and $r^2=x^2+y^2$. The superscript P indicates that this is the primary field. A horizontal dipole pointing in the y direction has the following components:

$$H_x^P=\frac{m_{Tx}}{4\pi}\frac{3xy}{R^5} \quad [39]$$

$$H_y^P=\frac{m_{Tx}}{4\pi}\left[\frac{3y^2}{R^5}-\frac{1}{R^3}\right] \quad [40]$$

and

$$H_z^P=\frac{m_{Tx}}{4\pi}\frac{3y(z-h)}{R^5} \quad [41]$$

Three of the most commonly used coil-pair geometries are shown in Figure 7 with the receiver coil located in the y–z plane ($x=0$) and the receiver and transmitter coils at the same height above the surface of the half-space ($z=h$). These geometries are horizontal coplanar (HCP), vertical coplanar (VCP), and vertical coaxial (VCX), where the horizontal and vertical orientations refer to the plane of the coil. The primary field for these measurement arrays can be computed using the appropriate expression for the dipole field components.

11.10.4.1.2 Mutual impedance

The concept of mutual impedance is a handy way to tie together the sensor voltage to the transmitter current. The voltage V_{Rx} in the receiver coil is related to the transmitter current I_{Tx} by the *mutual impedance* Z:

$$V_{Rx}=ZI_{Tx} \quad [42]$$

The receiver coil voltage is

$$V_{Rx}=-\mu_0 S_{Rx}\dot{H}_n \quad [43]$$

where $\dot{H}_n$ is the time derivative of the magnetic field normal to the plane of the receiver coil and S_{Rx} is the effective area of the receiver coil. When the magnetic field has $e^{i\omega t}$ time dependence, the receiver voltage can be expressed as

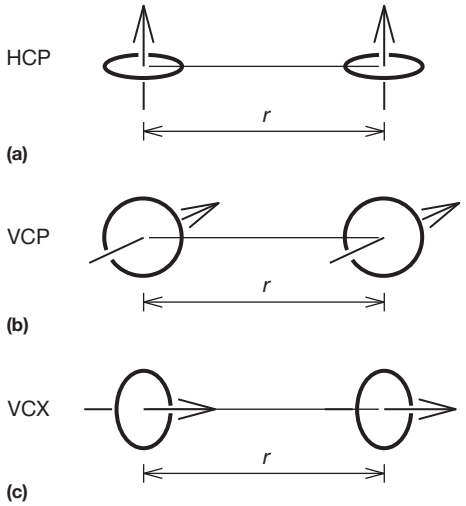


Figure 7 Geometry of common loop–loop electromagnetic systems. The geometries are described by the orientation of the plane of the loops. The dipole directions, which are orthogonal to the planes, are shown by the arrows. (a) Horizontal coplanar, HCP; (b) vertical coplanar, VCP; (c) vertical coaxial, VCX.

$$V_{Rx}=-i\omega\mu_0 S_{Rx}H_n \quad [44]$$

The magnetic field H_n or its time derivative $\dot{H}_n$ can be calculated for different currents and receiver moments from the mutual impedance by substituting eqn [44] or [43], respectively, into eqn [42]:

$$H_n=-\frac{ZI_{Tx}}{i\omega\mu_0 S_{Rx}} \quad [45]$$

$$\dot{H}_n=-\frac{ZI_{Tx}}{\mu_0 S_{Rx}} \quad [46]$$

As a means of normalizing the mutual impedance, we divide it by the free-space mutual impedance Z_0 to form the *mutual impedance ratio* Z/Z_0. The magnetic field is then given by

$$H_n(\omega)=-\frac{I_{Tx}}{i\omega\mu_0 S_{Rx}}(Z/Z_0)Z_0 \quad [47]$$

and the time derivative of the magnetic field by

$$\dot{H}_n(\omega)=-\frac{I_{Tx}}{\mu_0 S_{Rx}}(Z/Z_0)Z_0 \quad [48]$$

For HCP and VCP geometries, Z_0 is based on H_z^P in eqn [38] evaluated at $z=h$. Substituting H_z^P for H_n into eqn [45] and solving for Z_0 give

$$Z_0=\frac{i\omega\mu_0}{4\pi}\frac{S_{Rx}S_{Tx}}{r^3} \quad [49]$$

Similarly for the coaxial geometry, using H_y^P from eqn [40] evaluated at $x=0$, $y=r$, and $z=h$ gives

$$Z_0=-\frac{i\omega\mu_0}{2\pi}\frac{S_{Rx}S_{Tx}}{r^3} \quad [50]$$

which is -2 times the coplanar result. Note that the mutual impedance ratio of free space is 1.

Introducing eqn [49] into eqn [48], the time derivative of the magnetic field is given in terms of the mutual impedance ratio for the HCP and VCP geometries as

$$\dot{H}_{\mathrm{HCP/VCP}}(\omega) = -\frac{\mathrm{i}\omega S_{\mathrm{Tx}} I_{\mathrm{Tx}}}{4\pi r^3}(Z/Z_0)_{\mathrm{HCP/VCP}} \qquad [51]$$

and for the VCX geometry as

$$\dot{H}_{\mathrm{VCX}}(\omega) = \frac{\mathrm{i}\omega S_{\mathrm{Tx}} I_{\mathrm{Tx}}}{2\pi r^3}(Z/Z_0)_{\mathrm{VCX}} \qquad [52]$$

When the electric field is measured, eqn [42] takes on the form

$$V_{\mathrm{Rx}} = -E_{\mathrm{l}}\,\mathrm{d}l = ZI_{\mathrm{Tx}} \qquad [53]$$

where E_{l} is the component of electric field measured by a dipole of length dl. For analysis of the central-induction-loop, time-domain array, we will use the azimuthal electric field produced by a vertical magnetic dipole. The free-space mutual impedance used for normalization in this case is (Frischknecht, 1967)

$$Z_0 = -\frac{\mathrm{i}\omega\mu_0 S_{\mathrm{Tx}}\,\mathrm{d}l}{4\pi r^2} \qquad [54]$$

The azimuthal electric field due to a vertical magnetic dipole in terms of the mutual impedance ratio is given by

$$E_\phi(\omega) = \frac{\mathrm{i}\omega\mu_0 S_{\mathrm{Tx}} I_{\mathrm{Tx}}}{4\pi r^2}(Z/Z_0)_{\mathrm{WIRE}} \qquad [55]$$

11.10.4.1.3 Frequency-domain homogenous half-space response

Analytic expressions exist for the mutual impedance ratio when the transmitter and receiver coils are located on the surface of a half-space (Fitterman and Labson, 2005; Frischknecht, 1967; Keller and Frischknecht, 1966; Wait, 1955, 1982; Ward and Hohmann, 1988). For the three coil geometries shown in Figure 7, the mutual impedance ratios are as follows:

- Horizontal coplanar (HCP):

$$(Z/Z_0)_{\mathrm{HCP}} = \frac{2}{\gamma^2 r^2}\left[9 - \left(9 + 9\gamma r + 4\gamma^2 r^2 + \gamma^3 r^3\right)\mathrm{e}^{-\gamma r}\right] \qquad [56]$$

- Vertical coplanar (VCP):

$$(Z/Z_0)_{\mathrm{VCP}} = 2\left[1 - \frac{3}{\gamma^2 r^2} + \left(3 + 3\gamma r + \gamma^2 r^2\right)\frac{\mathrm{e}^{-\gamma r}}{\gamma^2 r^2}\right] \qquad [57]$$

- Vertical coaxial (VCX):

$$(Z/Z_0)_{\mathrm{VCX}} = \left[12 + 12\gamma r + 5\gamma^2 r^2 + \gamma^3 r^3\right]\frac{\mathrm{e}^{-\gamma r}}{\gamma^2 r^2} + 2 - \frac{12}{\gamma^2 r^2} \qquad [58]$$

where $\gamma^2 r^2 = \mathrm{i}\omega\mu_0\sigma r^2$

11.10.4.1.4 Induction number and skin depth

The mutual impedance ratios for a homogeneous half-space response can easily be written in terms of the *induction number* B, which is a dimensionless parameter that characterizes the induction process. The induction number is defined as

$$B = r/\delta \qquad [59]$$

where

$$\delta = \sqrt{\frac{2}{\omega\mu_0\sigma}} \qquad [60]$$

is the *skin depth*. It then follows that $\gamma r = (1+\mathrm{i})B$ and $\gamma^2 r^2 = \mathrm{i}\omega\mu_0\sigma r^2 = 2\mathrm{i}B^2$.

Under quasi-static conditions, Maxwell's equations describe a diffusive process that is mathematically the same as heat conduction in solids. Electromagnetic energy in the form of a plane wave impinging on a half-space will decrease inside the half-space with $\mathrm{e}^{-z/\delta}$ dependence, where z is the depth below the surface. Because the skin depth is a function of frequency, the depth of exploration is frequency-dependent with lower frequencies probing deeper. For non-plane-wave sources, such as magnetic dipoles, the depth of exploration also depends upon the intercoil separation and can be significantly less than the skin depth.

The induction process behaves in fundamentally different ways when the induction number is small ($B \ll 1$) compared with when it is large ($B \gg 1$). When B is small (the *inductive limit*), the process is controlled by the inductance of the ground. For large B (*resistive limit*), the process is controlled by the resistance of the ground.

Graphs of the inphase and quadrature components of the mutual impedance ratio for the HCP, VCP, and VCX coil geometries are shown in Figure 8. Two horizontal axes are shown: the lower one for frequency and the upper one for induction number. When the induction number is small ($B \ll 1$), the inphase and quadrature values approach 1 and 0, respectively, for all geometries. The inphase HCP response increases with increasing B reaching a maximum near B equal to 1.7 before decreasing to a minimum near B equal 6.0. At larger B, the response asymptotes at a value of 0. The quadrature HCP response starts and ends at a value of 0 with a minimum around B of 3.4. The VCP and VCX differ from the HCP response, but are similar to each other. The inphase VCP starts at 1 for small B and increases monotonically to 2 at large B. The inphase VCX behaves similarly except that the transition starts at a lower value of B, and there is a small maximum

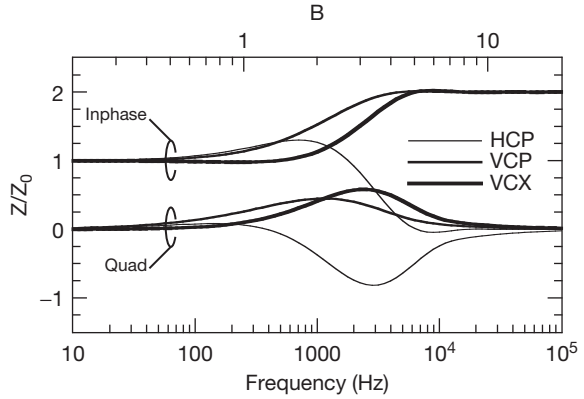


Figure 8 Mutual impedance ratio of three loop–loop configurations as a function of frequency. The inphase and quadrature curves have a value of 1 and 0, respectively, at low frequency. The corresponding induction number is shown on the upper axis. Half-space conductivity is $\sigma = 1/\pi^2 = 0.1013\,\mathrm{S\,m^{-1}}$. Coil separation is $r = 100\,\mathrm{m}$. (a) Inphase component; (b) quadrature component.

(1% of large B asymptote value) near B of 6.0. The quadrature component of the VCP and VCX responses increases at small B from a value of 0 to maxima near B of 2.1 and 3.2, respectively, before returning to 0 at large B.

The analytic expressions for the mutual impedance ratios in eqns [56] through [58] are only valid when the loops are on the surface of the half-space. When the loops are elevated above the half-space, closed form solutions do not exist; however, the analysis below for a layered half-space can be used to compute the response numerically.

11.10.4.1.5 Small-induction number behavior of homogeneous half-space

The small-induction number response of a half-space gives insight into the influence of model parameters. We start by making the substitution $x=\gamma r=(1+\mathrm{i})B$ in eqn [56] to obtain

$$\begin{aligned}(Z/Z_0)_{\mathrm{HCP}} &= \frac{2}{x^2}\left[9-\left(9+9x+4x^2+x^3\right)\mathrm{e}^{-x}\right] \\ &= \frac{2}{x^2}\left[9-\left(9+9x+4x^2+x^3\right)\sum_{n=0}^{\infty}(-1)^n\frac{x^n}{n!}\right]\end{aligned} \quad [61]$$

For small values of B, eqn [61] can be shown to equal (Fitterman and Labson, 2005; Kaufman and Keller, 1983)

$$(Z/Z_0)_{\mathrm{HCP}}^{\mathrm{LO}} = 1+\frac{\mathrm{i}B^2}{2} = 1+\frac{\mathrm{i}\omega\mu_0\sigma r^2}{4} \quad [62]$$

The inphase response is constant, while the quadrature response is proportional to conductivity and frequency. Similarly, the small-induction number response for the VCP and VCX geometries can be written as (Fitterman and Labson, 2005; Kaufman and Keller, 1983; Spies and Frischknecht, 1991; Wait, 1982)

$$(Z/Z_0)_{\mathrm{VCP}}^{\mathrm{LO}} = 1+\frac{\mathrm{i}B^2}{2} = 1+\frac{\mathrm{i}\omega\mu_0\sigma r^2}{4} \quad [63]$$

and

$$\begin{aligned}(Z/Z_0)_{\mathrm{VCX}}^{\mathrm{LO}} &= 1+\frac{(2\mathrm{i})^{3/2}}{15}B^3 = 1+\frac{(\mathrm{i}\omega\mu_0\sigma)^{3/2}}{15}r^3 \\ &= 1+\frac{-2+2\mathrm{i}}{15}B^3 = 1+\frac{-1+\mathrm{i}}{15\sqrt{2}}(\omega\mu_0\sigma)^{3/2}r^3\end{aligned} \quad [64]$$

The HCP and VCP results are the same, while the VCX response is different. In the low-induction number limit, the VCX quadrature component is proportional to $\sigma^{3/2}$, and the inphase component is 1 less an amount equal to the magnitude of the quadrature component. We will return to these low-induction number results when we discuss apparent resistivity functions.

11.10.4.1.6 Large-induction number behavior of homogeneous half-space

It is interesting to look at the large-induction number response for the HCP array, which is easily obtained from eqn [61] by letting x become large

$$(Z/Z_0)_{\mathrm{HCP}}^{\mathrm{HI}} = \frac{18}{x^2} = -\mathrm{i}\frac{9}{B^2} = -\mathrm{i}\frac{18}{\sigma\mu_0\omega r^2} \quad [65]$$

The asymptotic response is now purely imaginary and inversely proportional to conductivity and frequency. This behavior is very different from the low-induction number result and points out the nonlinear dependence of the electromagnetic response on the half-space conductivity over the complete range of induction numbers.

11.10.4.1.7 Frequency domain-layered half-space response

Many geologic situations can be modeled by a layered half-space. The electromagnetic response of this geometry has been described in a number of publications (Keller and Frischknecht, 1966; Wait, 1982; Ward and Hohmann, 1988). Using the notation of Wait (1982), the layered-earth mutual coupling ratios for the three geometries described earlier in the text are given by the following:

- Horizontal coplanar (HCP):

$$(Z/Z_0)_{\mathrm{HCP}} = 1+B^3T_0 \quad [66]$$

- Vertical coplanar (VCP):

$$(Z/Z_0)_{\mathrm{VCP}} = 1+B^2T_2 \quad [67]$$

- Vertical coaxial (VCX):

$$(Z/Z_0)_{\mathrm{VCX}} = 1+\frac{B^2}{2}(T_2-BT_0) \quad [68]$$

where the functions T_0 and T_2 are defined as

$$T_0 = \int_0^{\infty}\overline{R}(g)g^2\mathrm{e}^{-gA}J_0(gB)\mathrm{d}g \quad [69]$$

and

$$T_2 = \int_0^{\infty}\overline{R}(g)g\mathrm{e}^{-gA}J_1(gB)\mathrm{d}g \quad [70]$$

The function $\overline{R}(g)$ is given by the recursion relationship

$$\overline{R}(g) = \frac{Y_1-N_0}{Y_1+N_0} \quad [71]$$

where

$$Y_i = N_i\frac{Y_{i+1}+N_i\tanh\left[(g/2)^2+2\mathrm{i}\sigma_i/\sigma_1\right]^{1/2}D_i}{N_i+Y_{i+1}\tanh\left[(g/2)^2+2\mathrm{i}\sigma_i/\sigma_1\right]^{1/2}D_i} \quad [72]$$

$$Y_{M-1} = N_{M-1}\frac{N_M+N_{M-1}\tanh\left[(g/2)^2+2\mathrm{i}\sigma_{M-1}/\sigma_1\right]^{1/2}D_{M-1}}{N_{M-1}+N_M\tanh\left[(g/2)^2+2\mathrm{i}\sigma_{M-1}/\sigma_1\right]^{1/2}D_{M-1}} \quad [73]$$

$$N_i = \frac{\left(g^2+2\mathrm{i}\sigma_i/\sigma_1\right)^{1/2}}{\mathrm{i}\omega\mu_0\delta} \quad [74]$$

$$N_0 = \frac{g}{\mathrm{i}\omega\mu_0\delta} \quad [75]$$

and $A=(z+h)/\delta$ is the normalized sum of the transmitter and receiver coil heights, $B=r/\delta$ is the induction number, $D_i=2h_i/\delta$ is the normalized layer thickness, δ is the skin depth of the first layer, and J_0 and J_1 are the Bessel function of the first kind of orders zero and one, respectively.

The Hankel transforms in eqns [69] and [70] are evaluated numerically using a convolution method (Anderson, 1979; Johansen and Sorensen, 1979). These equations allow for the

coils to be above the surface of the half-space ($A>0$). The Hankel transforms as written previously in the text do not converge when both coils are located on the half-space ($A=0$). Convergent integrals are obtained by subtracting a term inside the integral corresponding to the secondary field over a homogeneous half-space and adding back the analytic expression for the homogenous half-space response outside the integral (Frischknecht, 1967).

Examples of two-layer responses that illustrate the basic behavior of a multilayer model are shown in Figures 9 and 10. The effect of the conductivity ratio σ_2/σ_1 is shown in Figure 9. Increasing σ_2 causes the maximum of the inphase and the minimum of the quadrature responses to move to a lower frequency, while the low- and high-frequency asymptotes remain the same as for a homogeneous half-space. In the limits of large and small induction numbers, the quadrature responses of all the models approach the same values. Also, the maximum that appears near $B=1$ for the homogeneous case ($\sigma_2/\sigma_1=1$) moves to a lower induction number as σ_2 increases. This suggests that at small induction numbers, the layered response is the same as for a half-space of conductivity σ_1, while at large induction numbers, the layered response behaves like a half-space of conductivity σ_2. These points will be discussed in more detail later.

Varying the first-layer thickness (see Figure 10) shifts the curves laterally and changes their shape. Again, the low- and high-frequency asymptotes are not changed. While the response is sensitive to model parameters, the curves do not intuitively indicate how conductivity varies with depth. This is an inherent problem with measured electromagnetic responses, particularly in the frequency domain. The standard approach is to transform the data to an apparent conductivity, which will be discussed later in the text.

11.10.4.1.8 Small-induction number behavior of layered half-space

Determining the small- and large-induction number asymptotes of a layered half-space is not for the faint of heart (Kaufman and Keller, 1983). Kaufman and Keller's analysis is exhaustive and leads to some very useful insights about the relationship between frequency-domain and time-domain methods. A simplified analysis, which serves our purposes, can be found in Wait (1982). The small-induction number behavior requires evaluation of the asymptotic behavior of the integrals T_0 and T_2 found in eqns [69] and [70], respectively. Wait's results are for the transmitter and receiver on the surface of the half-space ($h=0$ and $z=0$). The mutual impedance ratio for small induction number is given by

$$(Z/Z_0)^{\mathrm{LO}}_{\mathrm{geo}} = 1 + \frac{i\omega\mu_0 r^2}{4}\sum_{m=1}^{N}\sigma_m Q_m^{\mathrm{geo}} \qquad [76]$$

where Q_m^{geo} is a geometric factor specific to the coil geometry depending only on the layer thicknesses and intercoil spacing given by

$$Q_m^{\mathrm{HCP}} = \frac{1}{\left[1+4p_{m-1}^2\right]^{1/2}} - \frac{1}{\left[1+4p_m^2\right]^{1/2}} \qquad [77]$$

$$Q_m^{\mathrm{VCP}} = \sqrt{1+4p_{m-1}^2} - 2p_{m-1} - \sqrt{1+4p_m^2} + 2p_m \qquad [78]$$

and

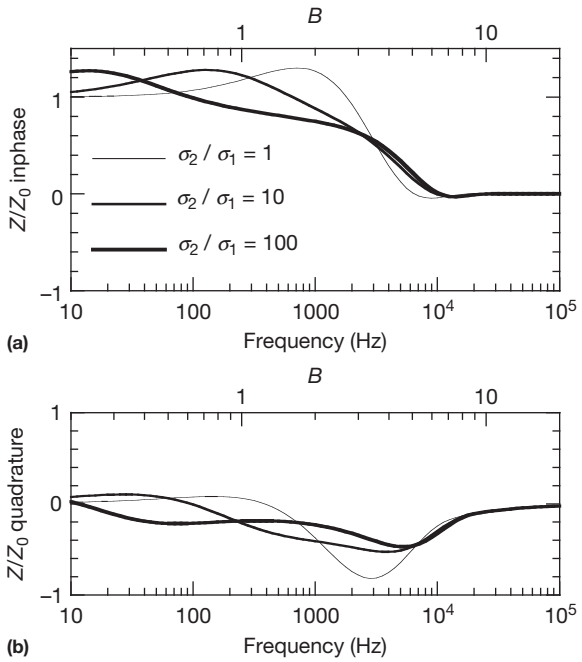


Figure 9 Mutual impedance ratio for HCP geometry over a two-layer model for different values of σ_2/σ_1 as a function of frequency and induction number. First-layer conductivity is $\sigma_1=1/\pi^2=0.1013\,\mathrm{S\,m^{-1}}$. h_1 is 25 m. Coil separation is 100 m. (a) Inphase component; (b) quadrature component.

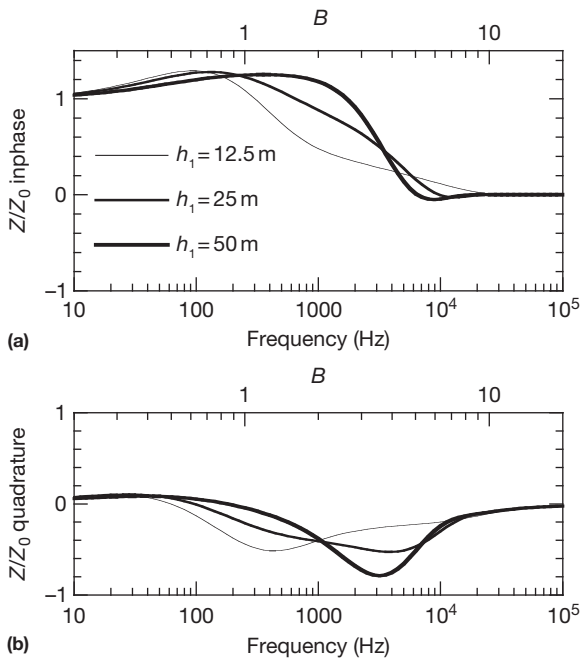


Figure 10 Mutual impedance ratio for HCP geometry over a two-layer model for different values of h_1 as a function of frequency and induction number. First-layer conductivity is $\sigma_1=1/\pi^2=0.1013\,\mathrm{S\,m^{-1}}$. σ_2/σ_1 is 10. Coil separation is 100 m. (a) Inphase component; (b) quadrature component.

$$Q_m^{\mathrm{VCX}} = \frac{\sqrt{1+4p_{m-1}^2}}{2} - p_{m-1} - \frac{1}{2\sqrt{1+4p_{m-1}^2}} - \frac{\sqrt{1+4p_m^2}}{2} + p_m + \frac{1}{2\sqrt{1+4p_m^2}} \quad [79]$$

where $p_m = d_m/r$ is the normalized depth to the bottom of the mth layer, $p_0 = 0$, and $p_N = \infty$. This is a remarkably simple result that arises because for small values of B, the currents flowing in different layers do not interact. That is to say, the induced current flowing in one layer does not induce a current in another layer. Furthermore, the secondary magnetic field is purely imaginary – entirely out-of-phase with respect to the primary magnetic field. When the low-induction number criterion is valid, the response can be rapidly computed without the need for numerical Hankel transforms. In practical terms, however, assuring that $B \ll 1$ is often not realizable. Nonetheless, the geometric factors give us insight about the vertical sensitivity of the various geometries.

Consider a two-layer model with an insulating second layer ($\sigma_2 = 0$). The model then represents a conductive slab of thickness d, and the geometric factors reduce to

$$Q^{\mathrm{HCP}} = 1 - \frac{1}{\sqrt{1+4p^2}} \quad [80]$$

$$Q^{\mathrm{VCP}} = 1 + 2p - \sqrt{1+4p^2} \quad [81]$$

and

$$Q^{\mathrm{VCX}} = -\frac{\sqrt{1+4p^2}}{2} + p + \frac{1}{2\sqrt{1+4p^2}} \quad [82]$$

where $p = d/r$. Plots of the geometric factors are given in **Figure 11 (a)**. The HCP geometric factor starts at 0 for small values of p and increases to a value of 1 as p increases. The cumulative response reaches a value of 0.5 at $p = 0.866$, meaning that half of the response is produced by the material between the surface and a depth equal to 86.6% of the intercoil separation. At $p = 1.936$, the cumulative response is 75%. The VCP cumulative response has a more rapid rise than the HCP response as p increases, meaning that the VCP array is more sensitive to near-surface material than the HCP array. The VCX geometric factor is peculiar in that it reaches a maximum of 0.15 at $p = 0.39$ before decreasing to zero as p increases further. Because of this behavior, the VCX array is not well suited for layered-earth geometry. In fact, this geometry is best suited for determining the dip of steeply dipping, conductive, planar features, such as fault zones (Frischknecht et al., 1991).

The results of the two-layer, insulating basement model can be used to investigate the depth sensitivity of the various geometries in the low-induction number regime. The sensitivity of an array to a particular depth can be determined by taking the derivative of Q^{geo} with respect to p. Differentiating eqns [80] through [82] results in

$$\frac{\mathrm{d}Q^{\mathrm{HCP}}}{\mathrm{d}p} = \frac{4p}{(1+4p^2)^{3/2}} \quad [83]$$

$$\frac{\mathrm{d}Q^{\mathrm{VCP}}}{\mathrm{d}p} = 2 - \frac{4p}{\sqrt{1+4p^2}} \quad [84]$$

and

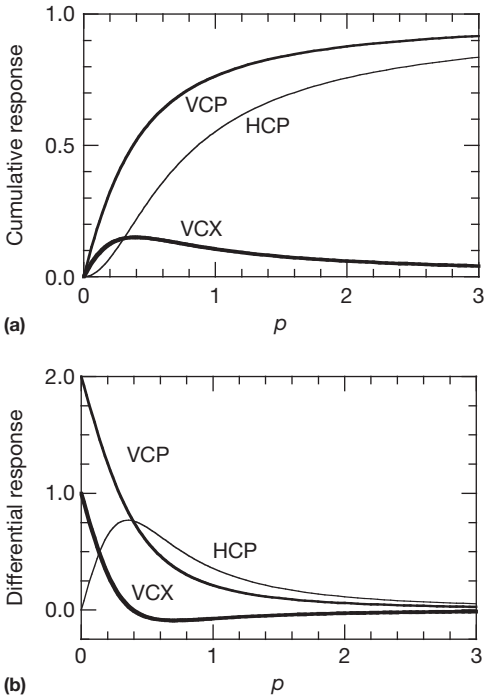


Figure 11 Response of various loop–loop geometries in the low-induction number regime as a function of normalized depth. (a) Cumulative response geometric factor Q^{geo}; (b) differential response geometric factor $\mathrm{d}Q^{\mathrm{geo}}/\mathrm{d}p$.

$$\frac{\mathrm{d}Q^{\mathrm{VCX}}}{\mathrm{d}p} = 1 - \frac{2p}{\sqrt{1+4p^2}} - \frac{2p}{(1+4p^2)^{3/2}} \quad [85]$$

Plots of these results are shown in **Figure 11(b)**, which confirm our earlier assessments of the relative depth of investigation of the various geometries based on the cumulative response function. The VCP array has the greatest sensitivity to near-surface material. The sensitivity drops to 10% of the surface value at $p = 1$. The HCP array, on the other hand, has zero differential response at the surface and reaches a maximum of 76% at a normalized depth of $p = 0.4$ before decreasing to zero at larger values of p. This gives the HCP geometry the greatest depth of exploration of the three arrays. The differential response of the VCX is 1 at the surface and becomes negative at $p = 0.4$. This is in agreement with the peculiar VCX cumulative response, which had a maximum at the same p value. As p increases further, the VCX differential response increases back to zero.

In these sensitivity calculations, it has been assumed that the coils are located on the surface of the half-space. In the VCP and VCX geometries, the coil centers are elevated by half a coil diameter. Elevating the coils above the half-space reduces the cumulative and differential sensitivities. This is not a large effect, provided the elevation of the coil centers is less than 1–3% of the intercoil spacing.

Instruments designed to work at low induction numbers are called *terrain conductivity meters*. These instruments are designed to read ground conductivity directly. This type of instrument is manufactured by Geonics (EM-34, EM-31, and EM-38), Geophex (GEM-2), Dualem (DUALEM-1, DUALEM-2, and DUALEM-4), and GF Instruments (CMD). In practice, these instruments often do not operate at a low

enough frequency to measure the true conductivity. We will address this issue later in the discussion of apparent resistivity.

11.10.4.1.9 Large-induction number behavior of layered half-space

Kaufman and Keller (1983) gave a detailed analysis of the large-induction number behavior of the layered half-space for a vertical magnetic dipole source. They showed that the response is controlled by the ratio of the total thickness H of layers above the basement to the skin depth of the first layer. When $H/\delta \gg 1$, the current does not penetrate very deeply, and the layers act as a half-space of resistivity ρ_1. For $H/\delta \ll 1$, the current penetrates to great depth and the stack of layers looks like a half-space of resistivity ρ_N. An overriding constraint is that the distance between the transmitter and receiver must be large compared to the skin depth $r/\delta \gg 1$, that is, the induction number is small; otherwise, current will flow primarily in the upper layers and a sounding cannot be made. An example of this sort of measurement is controlled-source audio-magnetotellurics (CSAMT) (Goldstein and Strangway, 1975).

11.10.4.1.10 Frequency-domain response of nonlayered geometries

Early frequency-domain exploration targets in the Canadian Shield were very conductive massive sulfide deposits in very resistive host material. This allows modeling the deposit as a conductive body in free space, which greatly simplifies the analysis. Keller and Frischknecht (1966) mentioned cases where the mathematical analyses for magnetic sources are available for simple geometries: a conducting sphere in an insulating whole space in the presence of a uniform (Wait, 1951) or dipolar (March, 1953; Wait, 1953) primary magnetic field and an infinitely conductive half-plane of vanishing thickness (finite conductance) (Wesley, 1958a,b). A thin conducting plate of finite dimensions is often a better geologic model than a half-plane model. Examples of how to numerically model finite geometry plates using a finite difference approach are given by Lamontagne and West (1971). Annan (1974) analyzed the same geometry using an eigencurrent formulation. Embedding the plate in a conducting half-space or dealing with more complicated geometries encountered in nature requires a different approach.

Starting in the 1970s, numerical modeling methods were developed for 2-D and 3-D geometries using finite difference and integral equation methods. Finite difference methods divide a half-space into discrete volumes of constant electrical property. The EM fields are modeled on a grid of nodes (or points), and the relationship between the fields at these nodes is given by Maxwell's equations with the derivatives replaced by difference operators. The grid must extend far enough from the 3-D body of interest so that effects caused by the edges of the grid are small. A large, sparse system of linear equations results that must be solved to obtain the field values. Integral equation methods avoid the problem of having to extend the computation grid to large distances; however, the system of linear equations that results is not sparse. A very good theoretical introduction to these techniques is found in Hohmann (1988), and a summary of more recent developments can be found in Oristaglio and Spies (1999). The development of 3-D EM modeling is an ongoing research area that will result in the routine use of 3-D techniques for data interpretation where appropriate.

Scale models have provided another approach to determining the response of 3-D geometries (Frischknecht, 1988; Keller and Frischknecht, 1966). Catalogs of curves from model studies were one of the principle ways of understanding electromagnetic profiling results over complex structures before the advent of 3-D numerical methods. Advances in instrumentation and materials available for model construction have made broadband frequency-domain and time-domain modeling possible. Early 3-D numerical results often challenged available computer resources, and models of high-conductivity contrasts between the target body and host material produced unreliable results. Model results played an important part in the development of 3-D numerical methods because they initially provided the only way to independently evaluate 3-D numerical results. As numerical methods improved and computer speed increased, the reliance on scale models has diminished. One area that has seen a significant resurgence in scale modeling is the characterization of the EM response of unexploded ordnance (UXO) (Asch et al., 2008; Wright et al., 2008).

11.10.4.2 Time-Domain Analysis

The EM systems described so far are characterized by a transmitter current that varies continuously; in particular, we have considered sinusoidally varying waveforms. Time-domain systems make use of current waveforms that are discontinuous, such as a step current turnoff:

$$I(t) = I_0[1 - u(t)] \qquad [86]$$

where $u(t)$ is the Heaviside step function. The time-domain response of this waveform can be computed by using Fourier analysis. The current waveform is Fourier-transformed into the frequency domain

$$I(\omega) = \int_{-\infty}^{\infty} I_0[1 - u(t)]e^{-i\omega t}dt = \pi\delta(\omega) - \frac{1}{i\omega} \qquad [87]$$

where $\delta(\omega)$ is the Dirac delta function. The voltage in the frequency domain is the mutual impedance for the geometry of interest $Z_{geo}(\omega)$ multiplied by the current:

$$V(\omega) = Z_{geo}(\omega)I(\omega) \qquad [88]$$

The inverse Fourier transform of eqn [88] gives the time-domain response:

$$V(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} V(\omega)e^{i\omega t}d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} Z_{geo}(\omega)I(\omega)e^{i\omega t}d\omega \qquad [89]$$

The most common time-domain geometry is the central-induction (CI) loop shown in **Figure 12(c)**. The array is composed of a circular transmitter loop of radius r through which a current I_{Tx} is passed. A receiver coil of area S_{Rx} is located at the center of the transmitter loop, and the voltage induced in the receiver coil V_{Rx} is recorded. This geometry can be derived from the azimuthal electric field $\mathbf{E}_\phi$ produced by a vertical magnetic dipole $\mathbf{m}_{Tx}$ (see **Figure 12(a)**). Integrating the electric field around the circumference of a loop of radius r gives the voltage measured by a voltmeter inserted into the loop circuit:

$$V_{CI} = 2\pi r E_\phi \qquad [90]$$

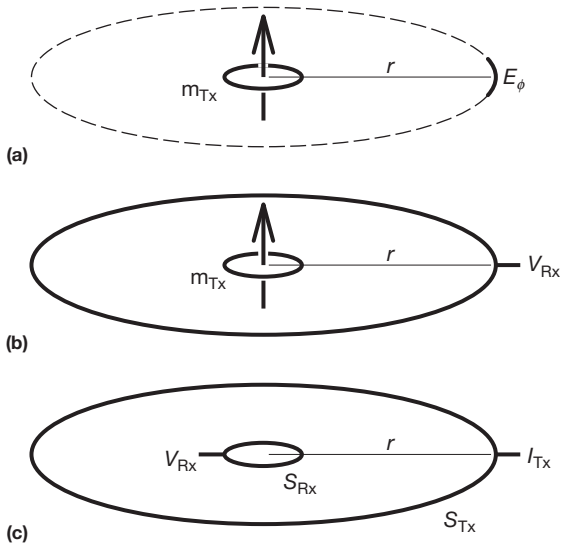


Figure 12 Transformation of azimuthal electric field produced by a vertical magnetic dipole source to central-induction-loop geometry. (a) Vertical magnetic transmitter dipole and azimuthal electric field; (b) integration of azimuthal electric field around outer loop; (c) interchange of transmitter and receiver.

where E_ϕ is the magnitude of the azimuthal electric field at distance r from the magnetic dipole.

11.10.4.2.1 Time-domain homogeneous half-space response

Using the expression for the azimuthal electric field from eqn [55] and the mutual impedance ratio for a vertical magnetic dipole and a grounded electric dipole (Fitterman and Labson, 2005; Frischknecht, 1967; Kaufman and Keller, 1983)

$$(Z/Z_0)_{\text{WIRE}} = \frac{2}{\gamma^2 r^2}\left[3 - \left(3 + 3\gamma r + \gamma^2 r^2\right)e^{-\gamma r}\right] \quad [91]$$

the electric field is written as

$$\begin{aligned} E_\phi(\omega) &= \frac{i\omega\mu_0 S_{Tx} I_{Tx}}{4\pi r^2}\frac{2}{\gamma^2 r^2}\left[3 - \left(3 + 3\gamma r + \gamma^2 r^2\right)e^{-\gamma r}\right] \\ &= \frac{S_{Tx} I_{Tx}}{2\pi r^4 \sigma}\left[3 - \left(3 + 3\gamma r + \gamma^2 r^2\right)e^{-\gamma r}\right] \end{aligned} \quad [92]$$

The inverse Fourier transform then gives the time-domain electric field:

$$\begin{aligned} E_\phi(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} E_\phi(\omega)\left[\pi\delta(\omega) - \frac{1}{i\omega}\right]e^{i\omega t}d\omega \\ &= \frac{S_{Tx} I_{Tx}}{2\pi r^4 \sigma}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[3 - \left(3 + 3\gamma r + \gamma^2 r^2\right)e^{-\gamma r}\right]\left[\pi\delta(\omega) - \frac{1}{i\omega}\right]e^{i\omega t}d\omega \end{aligned} \quad [93]$$

The contribution from the delta function integration is zero, leaving

$$E_\phi(t) = \frac{-S_{Tx} I_{Tx}}{2\pi r^4 \sigma}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[3 - \left(3 + 3\gamma r + \gamma^2 r^2\right)e^{-\gamma r}\right]\frac{e^{i\omega t}}{i\omega}d\omega \quad [94]$$

To evaluate this inverse Fourier transform, we make use of the following transform pairs:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{i\omega t}}{i\omega}d\omega = \begin{cases} 1, & t > 0 \\ 0, & t < 0 \end{cases} \quad [95]$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-\gamma r}}{i\omega}e^{i\omega t}d\omega = \text{erfc}\left(u/\sqrt{2}\right) = 1 - \Phi\left(u/\sqrt{2}\right) \quad [96]$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\gamma r e^{-\gamma r}}{i\omega}e^{i\omega t}d\omega = \sqrt{\frac{2}{\pi}}u e^{-u^2/2} \quad [97]$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\gamma^2 r^2 e^{-\gamma r}}{i\omega}e^{i\omega t}d\omega = \sqrt{\frac{2}{\pi}}u^3 e^{-u^2/2} \quad [98]$$

and

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\gamma^3 r^3 e^{-\gamma r}}{i\omega}e^{i\omega t}d\omega = \sqrt{\frac{2}{\pi}}u^3\left(u^2 - 1\right)e^{-u^2/2} \quad [99]$$

where

$$u = r\sqrt{\frac{\sigma\mu_0}{2t}} \quad [100]$$

and $\Phi(x) = \text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-s^2}ds$. (The transform in eqn [99] will be needed later for the HCP geometry.)

Applying the first three transforms to eqn [94] gives the voltage in the receiver dipole

$$E_\phi(t) = \frac{-3S_{Tx} I_{Tx}}{2\pi r^4 \sigma}\left[\Phi\left(u/\sqrt{2}\right) - \sqrt{\frac{2}{\pi}}u\left(1 + u^2/3\right)e^{-u^2/2}\right] \quad [101]$$

If we morph the receiver dipole into a loop of radius r with the transmitter dipole at it center, the voltage measured in this loop will be E_ϕ multiplied by the circumference of the loop:

$$V_{\text{LOOP}}(t) = \frac{-3S_{Tx} I_{Tx}}{r^3 \sigma}\left[\Phi\left(u/\sqrt{2}\right) - \sqrt{\frac{2}{\pi}}u\left(1 + u^2/3\right)e^{-u^2/2}\right] \quad [102]$$

Reciprocity allows interchanging the receiver and transmitter. In this case, we must replace the transmitter moment with the receiver coil moment to obtain the voltage measured by the central-induction-loop array:

$$V_{\text{CI}}(t) = \frac{-3S_{Rx} I_{Tx}}{r^3 \sigma}\left[\Phi\left(u/\sqrt{2}\right) - \sqrt{\frac{2}{\pi}}u\left(1 + u^2/3\right)e^{-u^2/2}\right] \quad [103]$$

Analysis of the time-domain HCP geometry proceeds in a similar manner. The time derivative of the vertical magnetic field in the frequency domain due to a vertical magnetic dipole is written using eqns [51] and [56]:

$$\begin{aligned} \dot{H}_{\text{HCP}}(\omega) &= -\frac{i\omega S_{Tx} I_{Tx}}{4\pi r^3}(Z/Z_0)_{\text{HCP}} \\ &= -\frac{i\omega S_{Tx} I_{Tx}}{4\pi r^3}\frac{2}{\gamma^2 r^2}\left[9 - \left(9 + 9\gamma r + 4\gamma^2 r^2 + \gamma^3 r^3\right)e^{-\gamma r}\right] \\ &= -\frac{S_{Tx} I_{Tx}}{2\pi\mu_0 r^5 \sigma}\left[9 - \left(9 + 9\gamma r + 4\gamma^2 r^2 + \gamma^3 r^3\right)e^{-\gamma r}\right] \end{aligned} \quad [104]$$

The time-domain result for a current step turnoff is given by

$$\dot{H}_{\text{HCP}}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\dot{H}_{\text{HCP}}(\omega)e^{i\omega t}}{-i\omega}d\omega \quad [105]$$

which can be evaluated using the inverse Fourier transforms in eqns [95] through [99] as

$$\dot{H}_{\mathrm{HCP}}(t)=\frac{S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi\mu_0 r^5\sigma}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[9-\left(9+9\gamma r+4\gamma^2r^2+\gamma^3r^3\right)\mathrm{e}^{-\gamma r}\right]\frac{\mathrm{e}^{\mathrm{i}\omega t}}{\mathrm{i}\omega}\mathrm{d}\omega$$
$$=\frac{9S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi\mu_0 r^5\sigma}\left[\Phi\left(u/\sqrt{2}\right)-\sqrt{\frac{2}{\pi}}u\left(1+u^2/3+u^4/9\right)\mathrm{e}^{-u^2/2}\right] \quad [106]$$

and the voltage induced in a receiver coil of moment S_{Rx} is

$$V_{\mathrm{HCP}}(t)=\frac{-9S_{\mathrm{Rx}}S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi r^5\sigma}\left[\Phi\left(u/\sqrt{2}\right)-\sqrt{\frac{2}{\pi}}u\left(1+u^2/3+u^4/9\right)\mathrm{e}^{-u^2/2}\right] \quad [107]$$

The normalized CI and HCP geometry results are plotted in Figure 13. The CI curve is initially constant and then decreases rapidly as time increases. The curve becomes a straight line on the log–log plot indicative of a power law dependence; it decays as $t^{-5/2}$. The CI curve does not change sign. The HCP curve starts out similar to the CI curve with a constant section followed by a rapid decrease in amplitude. In contract to the CI curve, the HCP curve goes through zero and reverses sign near the time of 2 ms. The curve then becomes similar to the CI curve and decays as $t^{-5/2}$. The behavior of the curve can be explained in terms of the diffusion of the induced currents in the half-space. When the transmitter current is turned off the induced current system initially mimics the current in the transmitter loop so as to maintain the magnetic field that existed just before the current was turned off. Immediately after turnoff, the induced current flows in a circular pattern directly under the transmitter loop, and very soon thereafter, the current flow starts to diffuse into the half-space. The location of the current maximum maintains a circular pattern; however, the maximum moves downward and outward, becoming more spread out as it diffuses. This pattern has been likened to a smoke ring diffusing into the ground (Nabighian, 1979). The HCP array behaves in a similar manner as the CI loop; only the transmitter loop is much smaller. For the HCP array, the current maximum is initially located between the transmitter and receiver coils. With increasing time, the maximum moves toward the receiver coil and then passes beyond it. As the maximum continues to move to the outside of the HCP array and the radius of the current system becomes large compared to the intercoil separation, the response becomes more and more like the CI response. At late enough time, the two responses behave identically indicating that the current systems are indistinguishable.

11.10.4.2.2 Dimensionless time-domain parameter

In a manner similar to the frequency-domain analysis, the time-domain response for a homogeneous half-space can be written in terms of a dimensionless parameter. Expressing the response as a dimensionless parameter makes it easy to scale the result to a different frequency (or time), intercoil separation, and/or conductivity. In the frequency domain, the dimensionless parameter is called the induction number B given in eqn [59]. The time domain has a corresponding dimensionless time parameter given by τ/r, where

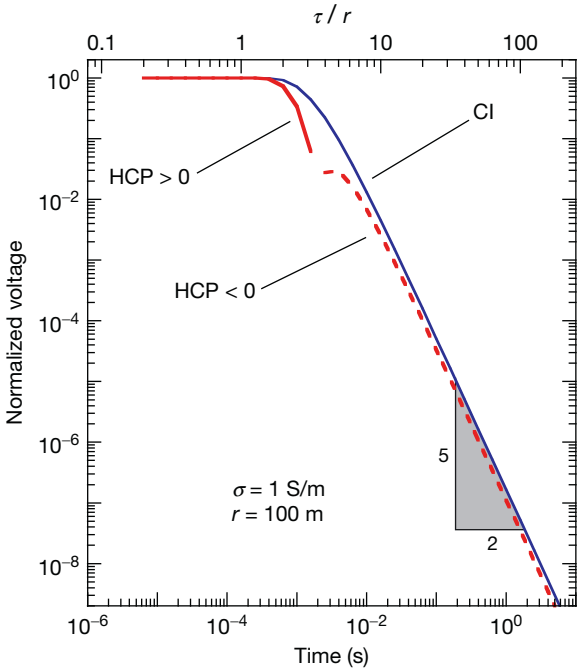


Figure 13 Normalized half-space time-domain response for central-induction (CI) loop and HCP arrays as a function of time and dimensionless time (τ/r). The CI curve is normalized by $-3S_{\mathrm{Tx}}I_{\mathrm{Tx}}/r^3\sigma$, and the HCP curve is normalized by $-9S_{\mathrm{Rx}}S_{\mathrm{Tx}}I_{\mathrm{Tx}}/2\pi r^5\sigma$. Half-space conductivity is 1 S m^{-1}. Loop radius is 100 m. Late-stage asymptotic time dependence ($t^{-5/2}$) is highlighted.

$$\tau=2\pi\sqrt{\frac{2t}{\mu_0\sigma}}=\sqrt{2\pi 10^7 t\rho} \quad [108]$$

is a scaled distance and r is the intercoil distance for the HCP array and the loop radius for the CI array. Another approach is to define a dimensionless parameter called the *diffusion depth*

$$\delta_{\mathrm{TD}}=\sqrt{\frac{2t}{\sigma\mu_0}} \quad [109]$$

which is analogous to the skin depth $\delta=\sqrt{2/\omega\mu_0\sigma}$. The diffusion depth is a measure of the extent of spreading of electromagnetic energy as it diffuses into the ground (Spies, 1989). The dimensionless time parameter can be written in terms of the diffusion depth as

$$\tau/r=\frac{2\pi\delta_{\mathrm{TD}}}{r}=\frac{1}{2\pi u} \quad [110]$$

The dimensionless time is also related to u as defined in eqn [100].

The time-domain response can be divided into different behaviors based on the value of τ/r. Small values of the parameter ($\tau/r\ll 1$) define the *early stage*, while large values ($\tau/r\gg 1$) define the *late stage* (Kaufman and Keller, 1983). The region between these limits is called the *intermediate stage*. We are careful not to refer to these different regions as early time or late time, a terminology that suggests a specific value of time, because the actual time when the various stages occurs depends on not only the time but also the half-space conductivity and the array size. The upper horizontal axis in Figure 13 shows the

τ/r value corresponding to the actual time based on the half-space conductivity and the Tx–Rx separation. A time-independent portion of the curve exists for $\tau/r < 1.5$; the rapid $t^{-5/2}$ decay starts around $\tau/r > 10$.

11.10.4.2.3 Early-stage and late-stage behavior of homogeneous half-space

The graph of the half-space time-domain response in Figure 13 reveals the time dependence of the early and late stages. Additional information about the stages can be determined by analysis of eqns [103] and [107] in the limit of small and large values of τ/r, which correspond to large and small values of u, respectively. The early-stage response of the CI array is given by

$$V_{\mathrm{CI}}^{\mathrm{ES}} = \lim_{u\to\infty} V_{\mathrm{CI}}(u) = \lim_{u\to\infty} \frac{-3S_{\mathrm{Rx}}I_{\mathrm{Tx}}}{r^3\sigma}\left[\Phi\left(u/\sqrt{2}\right) - \sqrt{\frac{2}{\pi}}u\left(1+u^2/3\right)e^{-u^2/2}\right] \quad [111]$$

Using $\lim_{x\to\infty}\Phi(x) = 1$ gives the result

$$V_{\mathrm{CI}}^{\mathrm{ES}} = \frac{-3S_{\mathrm{Rx}}I_{\mathrm{Tx}}}{r^3\sigma} \quad [112]$$

Similarly, the early-stage response for the HCP array can be obtained from eqn [107]:

$$V_{\mathrm{HCP}}^{\mathrm{ES}} = -\frac{9S_{\mathrm{Rx}}S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi r^5\sigma} \quad [113]$$

These results confirm the lack of time dependence seen in Figure 13; furthermore, they show that the voltage is inversely proportional to the conductivity σ; that is, the voltage is proportional to the resistivity $\rho = 1/\sigma$. The CI response is inversely proportional to r^3, and the HCP response is inversely proportional to r^5.

To determine the late-stage response, the error and exponential functions in eqns [103] and [106] are replaced with the following series expansions, respectively:

$$\Phi\left(u/\sqrt{2}\right) = \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n\left(u/\sqrt{2}\right)^{2n+1}}{n!(2n+1)} \quad (114)$$

and

$$e^{-u^2/2} = \sum_{n=0}^{\infty}\frac{(-1)^n(u^2/2)^n}{n!} \quad [115]$$

Applying these expressions to the CI array and gathering like terms give

$$V_{\mathrm{CI}}(t) = \frac{-3S_{\mathrm{Rx}}I_{\mathrm{Tx}}}{r^3\sigma} \times\sqrt{\frac{2}{\pi}}u\sum_{n=0}^{\infty}\frac{(-1)^n u^{2n}}{n!2^n}\left[\frac{1}{2n+1} - \left(1+\frac{u^2}{3}\right)\right] \quad [116]$$

For small values of u, eqn [116] becomes

$$V_{\mathrm{CI}}^{\mathrm{LS}}(t) = \frac{-S_{\mathrm{Rx}}I_{\mathrm{Tx}}}{5r^3\sigma}\sqrt{\frac{2}{\pi}}u^5 = \frac{-S_{\mathrm{Rx}}I_{\mathrm{Tx}}r^2}{20\sqrt{\pi}}\frac{\mu_0^{5/2}\sigma^{3/2}}{t^{5/2}} \quad [117]$$

For the HCP array, the corresponding series expression is

$$\dot{H}_{\mathrm{HCP}}(t) = \frac{9S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi r^5\mu_0\sigma} \sqrt{\frac{2}{\pi}}u\sum_{n=0}^{\infty}\frac{(-1)^n u^{2n}}{2^n n}\left[\frac{1}{2n+1} - \left(1+\frac{u^2}{3}+\frac{u^4}{9}\right)\right] \quad [118]$$

which for small u simplifies to

$$\dot{H}_{\mathrm{HCP}}^{\mathrm{LS}}(t) = -\frac{S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{5\pi\mu_0 r^5\sigma}\sqrt{\frac{2}{\pi}}u^5 = -\frac{S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{20\pi^{3/2}}\frac{\mu_0^{3/2}\sigma^{3/2}}{t^{5/2}} \quad [119]$$

and the voltage induced in the receiver coil is

$$V_{\mathrm{HCP}}^{\mathrm{LS}}(t) = \frac{S_{\mathrm{Rx}}S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{20\pi^{3/2}}\frac{\mu_0^{5/2}\sigma^{3/2}}{t^{5/2}} \quad [120]$$

The CI and HCP late-stage asymptotic expressions given in eqns [117] and [120], respectively, have very different time and conductivity dependences from the early-stage results discussed before. In the late stage, the voltage decays as $t^{-5/2}$, as was shown in Figure 13. The voltages varies as $\sigma^{3/2}$, a feature not apparent from the graph. This is in marked contrast to the early-stage result, which is proportional to the resistivity ($\rho = 1/\sigma$), suggesting that in the late stage, time-domain measurements are very good at detecting conductive targets. Not surprisingly, the CI and HCP late-stage results are nearly identical. Indeed, if the transmitter area in $\dot{H}_{\mathrm{HCP}}^{\mathrm{LS}}$ is replaced with an area equal to that of the CI array ($S_{\mathrm{Tx}} = \pi r^2$), the results are the same, indicating that in the late stage, the current distributions of the two arrays are equivalent. That is, at late enough time, the induced current due to the dipolar source diffuses and looks as though it was produced by a transmitter loop of radius r.

11.10.4.2.4 Time-domain-layered half-space response

The layered half-space time-domain response can be obtained by a Fourier transform approach similar to that used for the homogeneous half-space. We first consider the CI array, which requires calculation of the azimuthal electric field as the inverse Fourier transform of the product of $E_\phi(\omega)$ and the Fourier transform of the current waveform time dependence $1-u(t)$:

$$E_\phi(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}E_\phi(\omega)\left[\pi\delta(\omega) - \frac{1}{i\omega}\right]e^{i\omega t}d\omega = \frac{1}{2}E_\phi(0) + \frac{i}{2\pi}\int_{-\infty}^{\infty}\frac{E_\phi(\omega)}{\omega}e^{i\omega t}d\omega \quad [121]$$

As the primary electric field is zero, $E_\phi(0) = 0$. Because $E_\phi(t)$ is a real quantity, the imaginary part of the integral must equal zero:

$$0 = \frac{1}{2\pi}\mathrm{Im}\int_{-\infty}^{\infty}\frac{iE_\phi(\omega)}{\omega}e^{i\omega t}d\omega = \frac{1}{2\pi}\mathrm{Re}\int_{-\infty}^{\infty}\frac{E_\phi(\omega)}{\omega}e^{i\omega t}d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\mathrm{Re}E_\phi(\omega)\cos\omega t - \mathrm{Im}E_\phi(\omega)\sin\omega t}{\omega}d\omega \quad [122]$$

This condition is met if $\mathrm{Re}E_\phi(\omega) = \mathrm{Re}E_\phi(-\omega)$ and $\mathrm{Im}E_\phi(\omega) = -\mathrm{Im}E_\phi(-\omega)$. Applying these even and odd symmetry conditions to eqn [121] gives

$$E_\phi(t)=\frac{1}{2\pi}\mathrm{Re}\int_{-\infty}^{\infty}\frac{\mathrm{i}E_\phi(\omega)}{\omega}\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}\omega$$
$$=-\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\mathrm{Re}E_\phi(\omega)\sin\omega t+\mathrm{Im}E_\phi(\omega)\cos\omega t}{\omega}\mathrm{d}\omega \quad [123]$$
$$=-\frac{1}{\pi}\int_{0}^{\infty}\frac{\mathrm{Re}E_\phi(\omega)\sin\omega t+\mathrm{Im}E_\phi(\omega)\cos\omega t}{\omega}\mathrm{d}\omega$$

For $t<0$, $E_\phi(t)=0$ and eqn [123] become

$$0=\frac{1}{\pi}\int_{0}^{\infty}\frac{\mathrm{Re}E_\phi(\omega)\sin\omega t-\mathrm{Im}E_\phi(\omega)\cos\omega t}{\omega}\mathrm{d}\omega \quad (124)$$

Adding eqns [123] and [124] gives

$$E_\phi(t)=-\frac{2}{\pi}\int_{0}^{\infty}\frac{\mathrm{Im}E_\phi(\omega)\cos\omega t}{\omega}\mathrm{d}\omega \quad [125]$$

Alternatively, eqn [124] can be subtracted from eqn [123] to give

$$E_\phi(t)=-\frac{2}{\pi}\int_{0}^{\infty}\frac{\mathrm{Re}E_\phi(\omega)\sin\omega t}{\omega}\mathrm{d}\omega \quad [126]$$

The frequency-domain electric field is given in eqn [55], and the layered half-space mutual impedance ratio is given by Frischknecht (1967) as

$$(Z/Z_0)_{\mathrm{WIRE}}=1-B^2T_2 \quad [127]$$

where T_2 is defined in eqn [70]. Applying these substitutions to eqn [125], the electric field is

$$E_\phi(t)=-\frac{\mu_0 S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi^2r^2}\int_{0}^{\infty}\mathrm{Im}\left[\mathrm{i}(Z/Z_0)_{\mathrm{WIRE}}\right]\cos\omega t\,\mathrm{d}\omega$$
$$=-\frac{\mu_0 S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi^2r^2}\int_{0}^{\infty}\mathrm{Re}\left[(Z/Z_0)_{\mathrm{WIRE}}\right]\cos\omega t\,\mathrm{d}\omega \quad [128]$$

The CI array voltage is computed by multiplying eqn [128] by $2\pi r$ and replacing the transmitter area with the receiver coil area:

$$V_{\mathrm{CI}}(t)=-\frac{\mu_0 S_{\mathrm{Rx}}I_{\mathrm{Tx}}}{\pi r}\int_{0}^{\infty}\mathrm{Re}\left[(Z/Z_0)_{\mathrm{WIRE}}\right]\cos\omega t\,\mathrm{d}\omega \quad [129]$$

The HCP array is analyzed similarly with $\dot{H}_{\mathrm{HCP}}$ substituted for E_ϕ in eqn [125]

$$\dot{H}_{\mathrm{HCP}}(t)=-\frac{2}{\pi}\int_{0}^{\infty}\frac{\mathrm{Im}\dot{H}_{\mathrm{HCP}}(\omega)\cos\omega t}{\omega}\mathrm{d}\omega \quad [130]$$

Substituting eqn [51] for $\dot{H}_{\mathrm{HCP}}(\omega)$ gives

$$\dot{H}_{\mathrm{HCP}}(t)=\frac{S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi^2r^3}\int_{0}^{\infty}\frac{\mathrm{Re}(Z/Z_0)_{\mathrm{HCP}}\cos\omega t}{\omega}\mathrm{d}\omega \quad [131]$$

Calculation of the time-domain, layered media response requires two numerical transforms: first, a Hankel transform to compute the mutual impedance ratio in the frequency domain and, second, a cosine or sine transform to compute the time-dependent voltage. The latter transforms are done using a technique similar to the numerical Hankel transform (Anderson, 1975). The need for two transforms makes the time-domain calculation a bit slower than the frequency-domain calculation.

11.10.4.2.5 Theoretical basis for asymptotic behavior of layered half-space

Kaufman and Keller (1983) derived the early-stage and late-stage responses of a layered half-space by representing the frequency-domain response as a series expansion, rearranging terms, and then transforming to the time domain. This laborious procedure furnishes a simple result. The early-stage result for a layered half-space is the same as that for a homogeneous half-space with the half-space conductivity replaced by σ_1, that is, the first-layer conductivity. Recalling eqns [112] and [113], we write the early-stage responses as

$$V_{\mathrm{CI}}^{\mathrm{ES}}=\frac{-3S_{\mathrm{Rx}}I_{\mathrm{Tx}}}{r^3\sigma_1} \quad [132]$$

and

$$V_{\mathrm{HCP}}^{\mathrm{ES}}=-\frac{9S_{\mathrm{Rx}}S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{2\pi r^5\sigma_1} \quad [133]$$

The reason for this result is that at early enough time, the preponderance of the current is in the first layer, so only σ_1 has any influence on the response. By a similar argument, one would expect that at late enough time, most of the current is in the last layer and the response is controlled by σ_N. Indeed, this is the result obtained by Kaufman and Keller. From eqns [117], [119], and [120], the layered-earth late-stage responses are written as

$$V_{\mathrm{CI}}^{\mathrm{LS}}(t)=\frac{-S_{\mathrm{Rx}}I_{\mathrm{Tx}}r^2}{20\sqrt{\pi}}\frac{\mu_0{}^{5/2}\sigma_N^{3/2}}{t^{5/2}} \quad [134]$$

$$\dot{H}_{\mathrm{HCP}}^{\mathrm{LS}}(t)=-\frac{S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{20\pi^{3/2}}\frac{\mu^{3/2}\sigma_N^{3/2}}{t^{5/2}} \quad [135]$$

and

$$V_{\mathrm{HCP}}^{\mathrm{LS}}(t)=\frac{S_{\mathrm{Rx}}S_{\mathrm{Tx}}I_{\mathrm{Tx}}}{20\pi^{3/2}}\frac{\mu_0^{5/2}\sigma_N^{3/2}}{t^{5/2}} \quad [136]$$

Again, the responses have a high degree of sensitivity to the basement conductivity ($\sigma_N^{3/2}$) and decay rapidly with time ($t^{-5/2}$).

11.10.4.2.6 Examples of layered half-space response

Several examples of the CI array response for a two-layer half-space are shown in **Figure 14**. The first set of curves shows the effect of varying the first-layer conductivity (**Figure 14(a)**). The voltages have been normalized for the transmitter and array parameters by dividing by $-3S_{\mathrm{Tx}}I_{\mathrm{Tx}}/r^3$. The 1 S m^{-1} curve is for a homogenous half-space. The first thing to note is that the curves have a typical early-stage response, that is, there is no time dependence immediately after current turnoff. The amplitude of the voltage during this period is inversely proportional to the first-layer conductivity. The length of time that the early-stage behavior lasts increases as the first layer becomes more

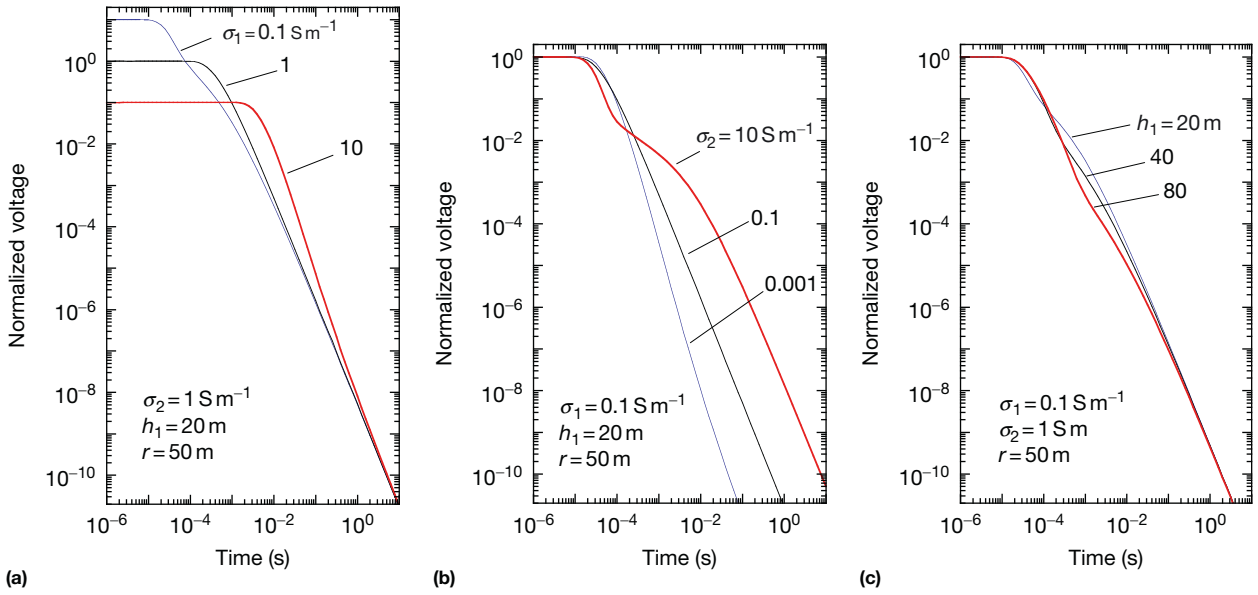


Figure 14 Normalized time-domain response for a CI loop over a two-layer half-space. Curves in panel (a) are normalized by $-3S_{Tx}I_{Tx}/r^3$ so that the early-stage values equal ρ_1. Curves in panels (b) and (c) are normalized by $3S_{Tx}I_{Tx}/r^3$. Loop radius is 50 m. Other model parameters are noted in figures. (a) First-layer conductivity σ_1 varies; (b) second-layer conductivity σ_2 varies; (c) first-layer thickness h_1 varies.

conductive because the higher conductivity slows down the diffusion of the current so that the conductivity of the underlying layer is not sensed until a later time. Following the early stage, the voltage starts to diminish, slowly at first and then much more rapidly with increasing time. Because the three models have the same basement conductivity, their responses approach that of the homogeneous half-space curve. The final portion of the curves shows the $t^{-5/2}$ behavior characteristic of the late stage. The time that it takes for the curves to attain the late-stage behavior and coalesce is proportional to the first-layer conductivity. The model with the more resistive first layer ($\sigma_1 = 0.1\,\mathrm{S\,m^{-1}}$) has an extra undulation during the transition between the early and late stages. This happens because the current diffuses rapidly through the resistive surface layer and is then slowed down by the more conductive second layer. Once most of the current has reached the second layer, the decay looks as if it were produced by a homogeneous half-space of conductivity σ_2. Notice that the ordering of the curves in the early stage, where smaller σ_1 corresponds to a larger voltage, inverts in the late stage where larger σ_2 corresponds to a larger voltage.

The second set of curves show the effect of varying the second-layer conductivity. Initially, the curves are the same because the first-layer conductivities are the same, and the current is predominantly in the first layer. As the current diffuses deeper, the influence of the second-layer conductivity is seen with the larger voltage corresponding to larger σ_2. The difference in σ_2 causes the curves to separate; however, they all have the same $t^{-5/2}$ behavior. The more resistive the second layer, the sooner the late stage is attained. The third set of curves (**Figure 14(c)**) shows the effect of varying the first-layer thickness for models that have a more conductive second layer. All curves look the same in the early and late stages, while each makes the transition between these stages in a different manner. As the first-layer thickness increases, the transition to the second layer is characterized by a more rapid decrease in voltage.

11.10.4.2.7 Time-domain response of nonlayered geometries

Nonlayered conductors behave in a very different manner from the layered-earth models described previously. The simplest example of this is a single-loop conductor with loop resistance R and self-inductance L. When excited by a magnetic dipole source, the voltage induced in a nearby receiver coil is given by the sum of the EMF due to the primary current in the transmitter dipole

$$V_{Rx}^{P}(t) = M_{Tx-Rx} I_{Tx} \delta(t) \qquad [137]$$

and the EMF due to the secondary current in the target loop

$$V_{Rx}^{S}(t) = \frac{M_{Tx-loop} M_{loop-Rx}}{L} I_0 \left[\frac{e^{-t/\tau} u(t)}{\tau} - \delta(t) \right] \qquad [138]$$

where M_{Tx-Rx}, $M_{Tx-loop}$, and $M_{loop-Rx}$ are the mutual inductances between the transmitter, receiver, and target loops; I_0 is the transmitter current; $\tau = L/R$ is the target-loop time constant; $u(t)$ is the Heaviside step function; and $\delta(t)$ is the Dirac delta function (Fitterman and Labson, 2005; Grant and West, 1965; West and Macnae, 1991). (NB: These expressions are for a current turnoff, while Grant and West's and West and Macnae's results are for a current turn-on.) The secondary voltage decays as an exponential whose time constant τ is given by the target parameters. This form, which is fundamentally different from a layered half-space that decays as a power law in the late stage, is characteristic of confined conductors (Kaufman, 1978). Large conductors in free space can be modeled by generalizing this technique so that the conductor is represented by an assemblage of loops that represent the eigencurrents that flow in the conductor (West and Macnae, 1991).

11.10.5 Apparent Resistivity

Geophysical electromagnetic instruments typically measure a voltage that is induced in a receiver coil. The measured voltage, which is expressed as magnitude and phase or inphase and quadrature component, is not intuitively related to the subsurface resistivity distribution, which the electromagnetic survey is trying to determine. Comparing measurements from one location to another is difficult and problematic, especially if any of the measurement parameters, which include things like transmitter current, transmitter loop size, transmitter–receiver separation, receiver gain, receiver coil moment, and frequency or time, have changed. Clearly, some sort of normalization is needed.

Apparent resistivity provides a way of affecting this normalization. While only a simple data transformation is usually involved, it is, nonetheless, a first step toward interpretation. Apparent resistivity is defined as the resistivity of a homogenous half-space that gives the same response as the actual measurement for the same measurement parameters. The resistivity is called apparent because it usually is not equal to the actual resistivity of the subsurface. The relationship between the apparent resistivity and the actual ground resistivity, even for a homogenous half-space, is complex for EM methods because the apparent resistivity usually has a nonlinear dependence on the actual resistivity. In addition, changes in measurement parameters often produce changes in apparent resistivity. As a consequence, a homogenous half-space may not appear to be homogenous as the measurement parameter (frequency, time, or intercoil separation) is varied. Nonetheless, transformation of a measured quantity to apparent resistivity is helpful and allows easier comparison of results. Spies and Eggers (1986) provided a detailed discussion of the concept of apparent resistivity.

11.10.5.1 General Concept

The idea of apparent resistivity was first developed for DC resistivity measurements by Wenner (1912), who relied upon the idea that a resistance or impedance could be converted to a resistivity by multiplying by a geometric factor K_g:

$$\rho_a = K_g Z = K_g \frac{V}{I} \qquad [139]$$

For the DC resistivity Wenner array (four equally spaced, collinear electrodes), the geometric factor is $2\pi a$, where a is the interelectrode spacing (Keller and Frischknecht, 1966). The transformation from impedance to resistivity is straightforward for DC resistivity because the impedance is a linear function of half-space resistivity. Determining the geometric factor for electromagnetic measurements requires a generalization of Wenner's approach.

Consider an electromagnetic measurement that is made over a homogenous half-space. Let the measured quantity be

$$V_h = F(\mathbf{p}, \rho_h) \qquad [140]$$

where $\mathbf{p}$ is the set of measurement parameters and ρ_h is the half-space resistivity. As we saw previously in the expressions for the frequency-domain and time-domain responses of a half-space, the function $F(\mathbf{p}, \rho_h)$ is a nonlinear function of the resistivity or conductivity. However, in the limits of small and large induction numbers, or early and late stages, the response can be written as the product of two functions

$$V_h = P(\mathbf{p})R(\rho_h) \qquad [141]$$

where $P(\mathbf{p})$ is a function of the measurement parameters and $R(\rho_h)$ is a function of the half-space resistivity. Solving eqn [141] for the resistivity function gives

$$R(\rho_h) = V_h / P(\mathbf{p}) \qquad [142]$$

When $R(\rho_h)$ has a single valued inverse, the half-space resistivity can be expressed as

$$\rho_h = R^{-1}[V_h / P(\mathbf{p})] \qquad [143]$$

Replacing V_h with the measured response V_m over the actual ground in eqn [143] gives the apparent resistivity

$$\rho_a = R^{-1}[V_m / P(\mathbf{p})] \qquad [144]$$

It is obvious looking at the homogeneous half-space expressions for the mutual impedance ratio in the frequency-domain responses, which are functions of $\gamma = i\omega\mu_0\sigma r$, and time-domain responses, which are functions of $u = r\sqrt{\sigma\mu_0/2t}$, that the factorization required by eqn [141] is not possible. Fortunately, we can make use of the asymptotic behavior in the limits of small and large induction numbers and early and late stages to solve this dilemma.

11.10.5.2 Frequency-Domain Apparent Conductivity Function

Consider the low-induction number HCP response given in eqn [62]. The imaginary part of this expression is

$$\mathrm{Im}(Z/Z_0)^{LO}_{HCP} = \frac{\omega\mu_0\sigma r^2}{4} \qquad [145]$$

Solving for the conductivity gives

$$\sigma = \frac{4}{\omega\mu_0 r^2}\mathrm{Im}(Z/Z_0)^{LO}_{HCP} \qquad [146]$$

If the mutual coupling ratio is replaced with the measured value, the conductivity is an apparent value:

$$\sigma_a^{LO} = \frac{4}{\omega\mu_0 r^2}\mathrm{Im}(Z/Z_0)_{HCP} \qquad [147]$$

Then, the low-induction number condition ($B \ll 1$) is valid; the conductivity is the true conductivity for a homogeneous half-space. Equation [147] is also valid for the VCP array because the low-induction number asymptote is the same for both arrays. In a similar fashion, a low-induction number apparent conductivity can be obtained for the VCX array starting with eqn [64]:

$$\begin{aligned}\mathrm{Im}(Z/Z_0)^{LO}_{VCX} &= \frac{\mathrm{Im}(i\omega\mu_0\sigma)^{3/2}}{15}r^3 = \frac{\mathrm{Im}\,e^{i3\pi/4}}{15}(\omega\mu_0\sigma)^{3/2}r^3 \\ &= \frac{\sqrt{2}}{30}(\omega\mu_0\sigma)^{3/2}r^3\end{aligned} \qquad [148]$$

Solving for the conductivity and substituting the actual impedance ratio give

$$\sigma_a^{LO} = \frac{1}{\omega\mu_0 r^2}\left[15\sqrt{2}\,\mathrm{Im}(Z/Z_0)_{VCX}\right]^{2/3} \qquad [149]$$

The behavior of these low-induction number apparent conductivity functions is shown in Figure 15. The upper panel shows the quadrature component of the mutual impedance ratio as a function of frequency and induction number. When plotted with a linear vertical axis, the numbers appear to be zero and show little change for frequencies less than 1 kHz corresponding to a induction number of about 0.3. At larger induction numbers, the response of all geometries starts to increase: The HCP curve increases very little, while the VCP and VCX curves increase more rapidly and attain a much higher value before decreasing. Converting these values to apparent conductivity produces the curves seen in the lower panel. At the lowest induction number (0.03), the apparent conductivity values are close to the half-space value of 250 mS m^{-1}. As the frequency and induction number increase, the apparent conductivity slowly decreases from the true value. The decrease continues, and the values eventually become negative, a physical impossibility. Remember, however, that these are only apparent conductivities. The HCP curve departs from the true value at the lowest frequency followed by the VCP and VCX curves in turn at higher frequencies. This downward bias in the apparent conductivity does not pose a problem unless you are trying to estimate the true conductivity of the ground. In that case, the behavior of the apparent conductivity must be considered.

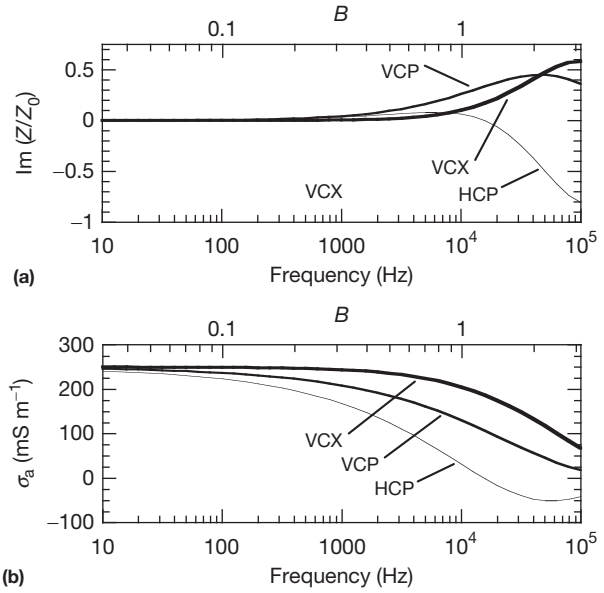


Figure 15 Quadrature component (a) and low-induction number apparent conductivity (b) as a function of frequency and induction number for HCP, VCP, and VCX geometries over a homogeneous half-space. The half-space conductivity is 0.25 S m^{-1}. The intercoil separation is 10 m.

There is a class of EM instruments called *terrain conductivity meters*, such as the Geonics EM-38, that are often used to determine the actual ground conductivity for agricultural monitoring and other purposes. The goal of the instrumentation is to easily determine ground conductivity without any need for interpretation. Figure 16(a) shows the apparent conductivity as a function of the true half-space conductivity for the EM-38, a handheld device with the coils mounted in a rigid fiberglass structure. Ideally, the graphs should be straight lines with a slope of 1 as shown by the dashed red line. The departure of the curves from the linear response suggests that determining the actual ground conductivity is not as simple as one would like. At low half-space conductivities, the apparent conductivity tracks fairly well; however, as the half-space becomes increasingly conductive, the apparent conductivity curves do not keep up. For the EM-38, the error in the conductivity is less than 10% for actual conductivities less than 160 mS m^{-1} (>6.25 Ω m) for the HCP geometry and less than 630 mS m^{-1} (>1.59 Ω m) for the VCP geometry. While the limit for the HCP geometry is low enough that it might pose a problem for some agricultural uses, the VCP limit would not. The upper horizontal scale in the graph

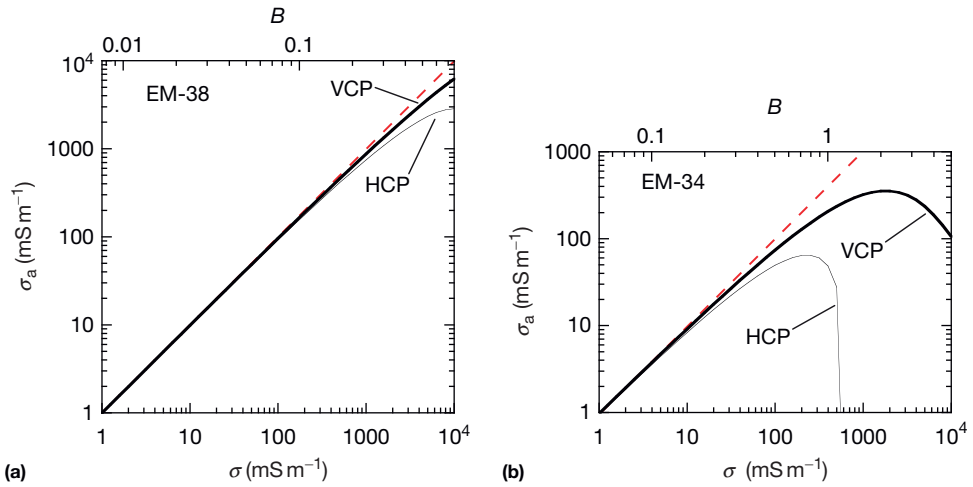


Figure 16 Terrain conductivity meter apparent conductivity σ_a over a homogeneous half-space of conductivity σ. The upper axis shows the corresponding induction number. HCP and VCP coil orientations are shown. The dashed red line shows a linear response. (a) Geonics EM-38: intercoil spacing is 1 m, and frequency is 14.5 kHz; (b) Geonics EM-34: intercoil spacing is 10 m, and frequency is 6.4 kHz.

shows the corresponding induction number based on the intercoil spacing (1 m) and frequency (14.5 kHz). The 10% error limits correspond to induction numbers of about 0.1 and 0.2, respectively, for the HCP and VCP geometries.

A similar set of curves is shown in **Figure 16(b)** for the Geonics EM-34, which has two 63-cm diameter coils that are connected by a cable. Intercoil separation can be set at 10, 20, or 40 m with a corresponding fixed frequency of 6400, 1600, or 400 Hz, respectively. Notice that these separation–frequency combinations maintain the same induction number for all three configurations ($B \sim r\sqrt{f}$); thus, the result in **Figure 16(b)** is valid for all EM-34 configurations. For the HCP geometry, the error in conductivity is less than 10% for half-space conductivities of less than 4 mS m^{-1} (>250 Ω m). For the VCP geometry, the error is less than 10% for conductivities of less than 16 mS m^{-1} (>62.5 Ω m). These are rather low conductivity values compared with those encountered in near-surface investigations. As the half-space conductivity increases, the departure from linear behavior becomes more pronounced leading to a maximum apparent conductivity value for each geometry and negative apparent conductivities in very conductive situations. This nonlinear behavior does not diminish the utility of the data, but rather requires that the user remember that the apparent conductivity value is an apparent, not an actual, value. The actual value can only be obtained by some sort of interpretation procedure. If we know for a fact that the ground was indeed a half-space, then curves like the ones shown in **Figure 16** could be used to determine the true conductivity. The Earth is rarely a homogenous half-space, so such a simple interpretational procedure is not realistic, and some sort of inversion procedure is often required.

The low-induction number apparent conductivity function is useful in a limited range of induction numbers. The same can be said for an apparent resistivity function based on the high-induction number behavior given in eqn [65]. Solving this equation for resistivity ($1/\sigma$), the following apparent resistivity function results in

$$\rho_a^{HI} = -\frac{\omega\mu_0 r^2}{18}\mathrm{Im}(Z/Z_0)_{HCP} \qquad (150)$$

This function is also useful in only a limited range of induction numbers. Wilt and Stark (1982) described a method for computing apparent resistivity that often works over a larger range of induction numbers. The response over a homogeneous half-space is computed for the measurement geometry as a function of induction number, for example, the HCP curves in **Figure 8**. One of the measured quantities, say $\mathrm{Im}(Z/Z_0)_{HCP}$, is entered on the ordinate, and the graph is used to find the corresponding abscissa, which gives an induction number. Equations [59] and [60] are then used to solve for the corresponding apparent conductivity. Because the induction number is a double-valued function of the measured quantity, there is an inherent ambiguity associated with in this method. Sometimes, the choice of induction number is obvious, while in other situations, it is not. Furthermore, there can exist situations where the measured response is outside the permissible range for a homogeneous half-space. In general, an apparent resistivity function that does not rely upon a low- or high-induction number asymptotic behavior can give ambiguous values.

Airborne FEM data is a case where single-frequency, apparent resistivity functions have proved useful in displaying data and understanding the underlying geology. Fraser (1978) described a procedure that is widely used for helicopter EM (HEM) data called the *pseudo-layer* apparent resistivity. HEM data depend upon the usual measurement array parameters plus the altitude of the coils above the ground. The coils are housed in a rigid structure called the *bird*, which is usually about 10-m long. HEM systems measure the helicopter altitude using a radar altimeter; however, the altitude data are often noisy due to the motion of the bird relative to the helicopter and the effects of tree canopy on the radar data. Fraser's approach is to add a high-resistivity layer, called the *pseudo-layer*, on top of the half-space to accommodate these measurement errors. Using the amplitude and phase of the HEM data, an apparent resistivity and bird height are found by matching the data to computed responses. Any difference between the calculated bird height and observed bird height is assigned to the pseudo-layer. The resulting apparent resistivity is unique.

Sengpiel (1988) formulated a way of using Fraser's pseudo-layer apparent resistivities to do an approximate inversion of HEM data. Each apparent resistivity is assigned a depth corresponding to the centroid of the current distribution that would be flowing in a homogenous half-space with a resistivity equal to the apparent resistivity. A spline is then fitted through the apparent resistivity-depth points to create a resistivity-depth function. Model studies have found that the resulting resistivity-depth function agrees rather well with the actual resistivity-depth function. Huang and Fraser (1996) modified Sengpiel's approach using an effective depth based on skin depth and the apparent depth to the top of the pseudo-layer half-space. Their method, called the differential apparent resistivity method, gives better results in many cases than Sengpiel's method. An example of differential resistivity HEM data analysis is given later in the text.

11.10.5.3 Time-Domain Apparent Resistivity Function

Time-domain apparent resistivity functions require the use of early- and late-stage asymptotic behavior to define a unique apparent resistivity. The most commonly used apparent resistivity function is the late-stage apparent resistivity. Starting with the late-stage voltage for the central-induction-loop array given in eqn [117], we solve for resistivity ($1/\sigma$):

$$\rho = \frac{\mu_0}{4\pi t}\left[\frac{2\pi r^2\mu_0 S_{Rx} I_{Tx}}{5t\left|V_{CI}^{LS}(t)\right|}\right]^{2/3} \qquad [151]$$

Replacing the late-stage voltage with the measured voltage gives an expression for the late-stage apparent resistivity:

$$\rho_{a\text{-}CI}^{LS} = \frac{\mu_0}{4\pi t}\left[\frac{2\pi r^2\mu_0 S_{Rx} I_{Tx}}{5t|V_{CI}|}\right]^{2/3} \qquad [152]$$

The HCP late-stage apparent resistivity function is obtained similarly starting with eqn [120]:

$$\rho_{a\text{-}HCP}^{LS} = \frac{\mu_0}{4\pi t}\left[\frac{2\mu_0 S_{Rx} S_{Tx} I_{Tx}}{5t|V_{HCP}|}\right]^{2/3} \qquad [153]$$

The early-stage asymptotic behavior also provides a way of defining apparent resistivity functions. In this case, we start with eqns [112] and [113] and solve for resistivity to obtain

$$\rho_{\text{a-CI}}^{\text{ES}} = \frac{r^3 |V_{\text{CI}}|}{3 S_{\text{Rx}} I_{\text{Tx}}} \quad [154]$$

and

$$\rho_{\text{a-HCP}}^{\text{ES}} = \frac{2\pi r^5 |V_{\text{HCP}}|}{9 S_{\text{Rx}} S_{\text{Tx}} I_{\text{Tx}}} \quad [155]$$

The early- and late-stage, central-induction, apparent resistivity functions are shown for a homogeneous half-space in Figure 17. The early-stage apparent resistivity is just a scaled version of the early-stage voltage curve. At small values of time corresponding to small τ/r, the curve is flat and has a value equal to the half-space resistivity of 1 Ω m. As time increases, the $\rho_{\text{a-CI}}^{\text{ES}}$ starts to drop below the half-space resistivity, slowly at first for $\tau/r < 2$, and then very rapidly. By the time $\tau/r > 5$, the apparent resistivity is decreasing as $t^{-5/2}$, the same rate as the decrease in the late-stage voltage. In contrast, the late-stage apparent resistivity has values much larger than the half-space resistivity for $\tau/r < 2$ and then starts to gradually approach the half-space resistivity. By $\tau/r > 10$, $\rho_{\text{a-CI}}^{\text{LS}}$ has settled at the half-space value of 1 Ω m. As pointed out previously, the actual time that corresponds to a dimensionless time depends upon the half-space resistivity and the transmitter loop radius. The important thing to remember is that the apparent resistivity is equal to the actual resistivity only in the early stage for $\rho_{\text{a-CI}}^{\text{ES}}$ and in the late stage for $\rho_{\text{a-CI}}^{\text{LS}}$. For time values corresponding to $\tau/r < 2$, $\rho_{\text{a-CI}}^{\text{LS}}$ will always be unrealistically high, and for time values corresponding to $\tau/r > 5$, $\rho_{\text{a-CI}}^{\text{ES}}$ will be unrealistically low. As with the frequency-domain apparent resistivity functions, these characteristics of the time-domain apparent resistivity function do not denigrate their utility, but rather, we must be aware of their behavior in order to avoid misinterpreting their meaning.

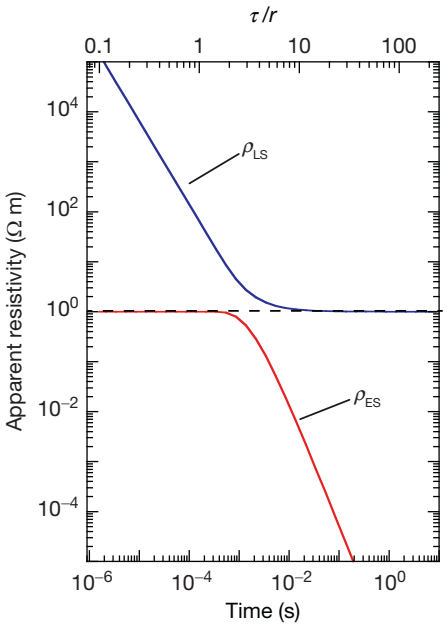


Figure 17 Early- and late-stage apparent resistivity functions for the central-induction-loop array over a homogeneous half-space. The half-space resistivity is 1 Ω m (dashed black line). The loop radius is 100 m. The upper axis shows the corresponding dimensionless time τ/r.

To further demonstrate the behavior of time-domain apparent resistivity curves, the voltage transients from Figure 14 have been transformed into late-stage apparent resistivity. Figure 18(a) shows a set of sounding curves with different first-layer resistivity values underlain by a fixed second-layer resistivity of 1 Ω m. One of the curves is for a homogenous half-space similar to the late-stage apparent resistivity curve shown in Figure 17. At an early time, all of the curves have apparent resistivities that far exceed the first-layer resistivity, which is in the range of 0.1–10 Ω m. For times less than 20 μs, the larger the first-layer resistivity, the smaller the apparent resistivity. At a time of 20 μs, the $\rho_1 = 10$ Ω m curve starts to decay less rapidly because current diffusion is starting to be controlled by the more conductive second layer and the 10 Ω m curve crosses the other two curves. At increasing time, 10 Ω m curve attains the resistivity of the second layer (1 Ω m), as predicted by the late-stage asymptotic behavior of a layered model. The lowest first-layer resistivity curve ($\rho_1 = 0.1\,\Omega$m) descends for a longer time than the other curves. It then goes through a minimum near 10 ms before increasing and approaching the basement resistivity of 1 Ω m. The value of the minimum is slightly greater than the first-layer resistivity. If the first layer were thicker, the minimum apparent resistivity would approach the true resistivity of the conductive first layer. For the $\rho_1 = 10\,\Omega$m curve, the change in

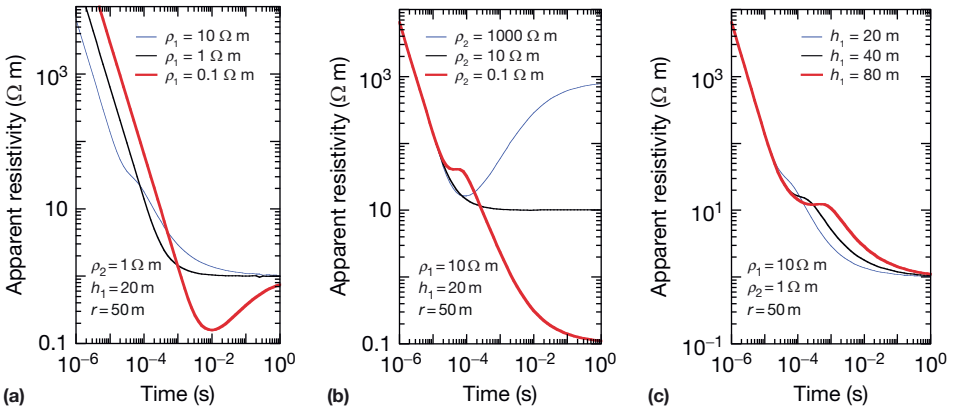


Figure 18 Late-stage apparent resistivity for CI loop over a two-layer half-space. Models are the same as in Figure 14. Loop radius is 50 m. Other model parameters are noted in figures. (a) First-layer resistivity ρ_1 varies; (b) second-layer resistivity ρ_2 varies; (c) first-layer thickness h_1 varies.

slope near 20 μs mentioned before is caused by the current diffusion being affected by the more conductive second layer. Notice that the ordering of the curves, which is inverse to the value of the first-layer resistivity at times less than 25 μs, is in the order of the first-layer resistivity at times greater than 1 ms. This reflects the transition from the early stage to the late stage.

The effect of changing the second-layer resistivity is illustrated in Figure 18(b). The first-layer resistivity is fixed at 10 Ω m for all curves, one of which is a homogeneous half-space. All three curves are similar at early time because current diffusion is controlled by the first-layer resistivity. The resistive basement curve exhibits a minimum whose value is greater than the first-layer resistivity. Had the first layer been thicker, the minimum apparent resistivity value would have been closer to the actual first-layer resistivity. The conductive second-layer model shows a small bump around 0.1 ms. This maximum is characteristic of a decrease in layer resistivity occurring in the early part of the transient and does not indicate the presence of a thin resistive layer in the model. In fact, if a thin resistive layer were sandwiched between the first and second layers in any of these models, its effect would be small to imperceptible due to equivalence. As time increases, the three curves separate as they each approach their respective second-layer resistivities.

Lastly, we look at the effect of changes in the first-layer thickness (Figure 18(c)). The first- and second-layer resistivities are fixed at 10 and 1 Ω m, respectively. Until a time of about 20 μs, the curves are the same because the current is being influenced only by the first-layer resistivity. At this time, the $h_1 = 20$ m curve starts to depart from the other two curves as the less-resistive second layer starts to influence the current flow. The $h_1 = 40$ m curve separates from the $h_1 = 80$ m curve at around 60 μs.

11.10.6 Interpretation

The goal of electromagnetic measurements is to obtain subsurface information to solve environmental, geotechnical, or exploration problems. As we have seen, raw electromagnetic data are difficult to understand. Conversion of these data to apparent resistivity (or conductivity) allows easier comparison of data from different locations. Apparent resistivity can also give a sense of how electrical properties vary with depth. While this is an improvement over the raw data, it usually is not adequate to answer quantitative questions, such as the following: What is the actual resistivity at a particular depth, or at what depth does the transition to a higher resistivity occur? Answering these questions requires a model of the subsurface resistivity distribution, which is the product of a geophysical interpretation. It is called a geophysical interpretation because it says nothing about the geologic nature of the subsurface. The geophysical model does not specify if a 10 Ω m layer found between depths of 22 and 37 m is a saltwater-saturated sandstone, an altered basalt, or a clay. The answer to this and similar questions is what makes up a geologic or hydrologic interpretation. The geophysical interpretation results in a model that can produce the measured electromagnetic response. It does not care what the nature of the material is as long as the measured and computed electromagnetic responses agree. Over the years, geophysical interpretations have been obtained by a variety of means including curve matching of data to precomputed curves, forward calculation of model response with manual adjustment of parameters to minimize the error between measured and computed responses, and inversion whereby a mathematical procedure is used to find the model that best fits the data subject to data uncertainties and specified constraints. The use of inversion has become standard for geophysical interpretation.

Geologic interpretation is the process of assigning geologic descriptions to the various components of the geophysical model. This is done using knowledge of the resistivity ranges of rock types such as granite, sandstone, or clay and anticipated resistivity transitions that can occur across geologic or hydrologic boundaries such as contacts, facies changes, faults, or saltwater fronts. Because the range of resistivity for a particular rock type is large and the resistivity ranges of various rock types often overlap as shown previously (see Figure 5), geologic interpretations based on resistivity alone are often nonunique. The use of auxiliary information such as the geology of the area from mapping, lithologic logs from well cuttings and cores, borehole geophysical logs, and field observations made while conducting the geophysical survey can reduce the nonuniqueness. There is also an interplay between the development of the geophysical and geologic models. Because of the inherent uncertainty in geophysical models, there are a multitude of models that can fit any particular set of observations, and geologic information often suggests alternative geophysical models that must be examined to see if they are allowed by the data. A further complication is that 'known' geology is sometimes specious, being based on inaccurate information or assumptions that are incorrect. Resolution of the incongruity between the geologic and geophysical models requires a give and take between the two realms. A satisfactory resolution is achieved when the revised model fits the data and the corresponding geologic story is meaningful.

11.10.6.1 Inversion

Inversion is a procedure for determining a model that produces the best agreement between the observed response d_j^{obs} and the model response $F_j[m]$, where F_j is the forward model operator and m is the model. The observations usually contain noise due to geologic inhomogeneities, electromagnetic noise, and incompatibilities between the model and actual geometry. The model contains information about the measurement geometry and the resistivity of the ground given as either a distribution, a set of layers, or other geometric features. Selection of the model from the infinity of possibilities is constrained by the data. It is not uncommon for the number of model parameters to exceed the number of observations. Clearly, additional information must be used to determine the model.

Oldenburg and Li (2005) described in detail a common approach to this problem. A *data objective function* is defined as

$$\phi_{\text{d}} = \left\| \mathbf{W}_{\text{d}} \left(\mathbf{d}^{\text{obs}} - \mathbf{F}[m] \right) \right\|^2 \qquad [156]$$

where $\mathbf{d}^{\text{obs}}$ is a vector of the observations, $\mathbf{F}[m]$ is a vector of the model responses, and $\mathbf{W}_{\text{d}}$ is a diagonal matrix whose elements are $1/\epsilon_j$, where ϵ_j is the standard deviation of a datum. Models that minimize ϕ_{d} attempt to honor the data. A *model objective function* is also defined as

$$\phi_m = \alpha_s \int (m - m_{ref})^2 ds + \alpha_x \int \left(\frac{dm}{dx}\right)^2 dx \qquad [157]$$

which puts constraints on the model. The first term favors models that are closer to some reference model m_{ref}, and the second term favors models that are smooth, that is, their spatial derivatives are small in absolute terms. Other terms can be added to the model objective function, such as ones that favor lateral consistency between adjacent models (Auken and Christiansen, 2004; Auken et al., 2005). The goal is to find a model that minimizes the model objective function and reproduces the data as well as possible. This is done by recasting the problem into an *unconstrained optimization* problem to minimize

$$\phi(m) = \phi_d + \beta\phi_m \qquad [158]$$

where β is a nonnegative constant called the regularization or trade-off parameter. The *regularization parameter* sets a balance between selecting a simple model without erratic, extraneous, or unrealistic resistivity variations and reproducing the data with an uncertainty consistent with the uncertainty in the observations.

As shown previously, the electrical resistivity of geologic materials spans a range of nearly six orders of magnitude (see **Figure 5**). Trying to determine a parameter with so large a range is very difficult. To address this issue, most electromagnetic inversions solve for the logarithm of both the conductivities and model dimensions (Hohmann and Raiche, 1988; Johansen, 1977). Bounds can also be put on model parameters to force the inversion to return physically realistic values. This can be done by adding a term to the model objective function that penalizes model parameters that are outside of a specified range. Oldenburg and Li (2005) described a logarithmic barrier method. Lagrange multipliers can also be used to enforce parameter bounds (Tarantola, 1987). The resulting inversion is termed an unconstrained optimization problem because the parameter constraints are part of the objective function.

The process of inversion is, in general, nonunique and unstable because the number of data points is finite, the data are noisy, and the physics of the problem may render the response insensitive to certain model parameters. In the electromagnetic literature, this last issue is called *equivalence*. For layered-earth geometry, equivalence can occur when a thin layer is present, that is, a layer whose thickness is small compared to its depth of burial. If the layer is conductive compared to the surrounding material, the response is sensitive only to the conductance $S = \sigma h = h/\rho$ of the thin layer, where h is the thickness of the thin layer and ρ is its resistivity. This is called *S equivalence*. On the other hand, if the thin layer is much more resistive than the surrounding material, the response is sensitive only to the layer thickness h and not its resistivity ρ. This is called *H equivalence*. In essence, the thin, high-resistivity layer behaves like free space so we cannot determine its resistivity. More accurate measurements will not resolve equivalence because it is inherent to the physics of the problem. The use of auxiliary data, as well as information from complementary geophysical methods, is the only way to reduce the uncertainty caused by equivalence.

Layered-earth (1-D) electromagnetic inversion schemes are often used in situations where the geology is laterally varying. If lateral changes are small, a quasi-2-D interpretation can be approximated by inverting the data from each measurement site and then stitching them together when they are displayed. An example of this is given in the succeeding text for airborne EM data.

Natural and man-made electromagnetic noise will also affect inversion results. Power-line noise is ubiquitous, though naturally occurring EM noise can be reduced by a variety of methods. In the frequency-domain, filtering is an effective means of dealing with it.

11.10.6.2 Noise in Electromagnetic Data

Several types of noise are found in electromagnetic data that can limit its usefulness. First among these is electromagnetic noise from natural or man-made sources. Natural electromagnetic noise is termed *sferics* and is produced by electric discharges associated with thunderstorms (McCracken et al., 1986). The spectrum of sferic noise, which can vary with time of day, latitude, and season, is broadband. In contrast, man-made electromagnetic noise, which is usually caused by current flowing in power lines or nearby high-power, commercial radio transmitters, has a band-limited spectrum. Frequency-domain instrumentation can use filters to notch out offending signals at the primary power-line frequency and its harmonics. Alternatively, passband filters can be used to allow only the frequencies used by the geophysical system to be measured. These approaches will not work for time-domain instrumentation, which needs to be broadband to accurately record electromagnetic transients. The easiest way to reduce power-line noise is to not work too close to power lines, particularly high-tension lines that carry very large currents. At distances of less than 100–400 m, power-line noise can be severe. Another approach to reduce the effect of noise is to increase the transmitter moment thereby increasing signal strength. Sometimes, this will not work because the power-line noise is so large that it overloads the instrumentation front-end. Data stacking can also be used to increase the signal-to-noise ratio (Macnae et al., 1984). Data stacking in the time domain requires that the single current turnoff that we used in our analyses be replaced with a repetitive signal – often a 50% duty-cycle square wave. Each current turn-on and current turnoff of this waveform have a transient associated with them. Usually, only the transient associated with the current turnoff is recorded. By choosing the repetition frequency of the current square wave to be an integral subharmonic of the power-line frequency (1/2, 1/3, etc.), constant-phase, power-line noise is reduced.

Electric fences used to control livestock are another source of electromagnetic noise. The electrified portion of the fence consists of one or more strands of wire that are insulated from the fence posts. A pulsating voltage source is attached between the wires and a grounding system. The voltage is typically 1–7 kV, and the pulse is 1 s on followed by 1 s off. The energized fences can be as much as 50 km long. The signals produced by these systems make working near them very difficult. The only reliable way to deal with them is to temporarily turn off the fence charger while geophysical measurements are being made.

Background EM noise limits the time after transmitter turnoff to which TEM measurements can be made. Fitterman

(1989) showed how a detection limit associated with natural source EM noise can be drawn on apparent resistivity-time plots. The expression for late-stage apparent resistivity from eqn [152] is rewritten as

$$\rho_a^{detc} = \frac{\mu_0}{4\pi t}\left[\frac{2\pi r^2 \mu_0 S_{Rx} I_{Tx}}{5t|V_{CI}|}\right]^{2/3} = \frac{1}{4\pi}\left[\frac{\mu_0}{t}\right]^{5/3}\left[\frac{M_{Tx}}{5N_{EM}}\right]^{2/3} \quad [159]$$

where the ratio of the sensor voltage V_{CI} to the receiver moment S_{Rx} has been replaced with electromagnetic noise level N_{EM}, which is typically 10^{-9} to 10^{-10} V m^{-2}, and the transmitter moment is $M_{Tx} = 2\pi r^2 I_{Tx}$ for a circular loop. This expression produces a line with a slope of $-5/3$ on a log–log plot. Detectable points will fall to the left of the line. An example of the use of this expression will be given later.

Metallic objects on or below the surface are potential electromagnetic noise sources. Objects that contain closed circuit paths such as steel plates, barrels, loops of wire, or steel fences can be energized by the primary EM field. The currents induced in these objects produce fields that may be detected by the receiver, adding noise to the measurement. Removal of these signals is often impossible, and the corrupted data cannot be used. Fortunately, the signals produced by metallic objects are usually peculiar enough that they can be recognized. If detected in the field, the offending object can sometimes be moved or the measurement point relocated away from the offending object. Buried steel pipes can distort the currents induced in the ground, producing very large signals. In the case of transient soundings, apparent resistivity-time data plot as long, linear features on log–log coordinates suggestive of a very conductive layer at depth. The tip-off to the presence of a pipe anomaly is that the apparent resistivities are too low to be of geologic origin.

11.10.7 Practical Aspects of Survey Design

At this point, you may be wondering how you are going to put all this theory to practical use. The best way to start is to forget about the theory and ask the question: What am I trying to determine about the subsurface? It also does not hurt to draw a picture or conceptual model of what you think the subsurface conditions are. Your picture should include things like the expected rock types, their geometry (layered, dipping, faulted, or folded), and the saturating pore fluid, if any. Some aspects of the diagram may be more reliable than others. For example, existing geologic mapping may have identified rock types that are likely to be present in the subsurface. Logs from a nearby drill hole can give direct information about subsurface geology and the depths of the various interfaces. Previous geophysical surveys can give additional information and possibly resistivity estimates of the expected geologic units. The features that you are certain about might be drawn with bolder lines than the features that are less reliable. Next, mark with question marks features that you want to determine geophysically. These might be the thickness of sedimentary fill over crystalline bedrock, whether or not a unit is freshwater- or saltwater-saturated, or the location of the contact between two geologic units. At this point, you should have a sketch of the situation with your objectives identified.

The next step is to associate numbers with features in your conceptual model. This will transform the sketch into a geophysical model. Write down the depths, thicknesses, and widths of the various units. If you do not know these values exactly, then use a likely range of values. Be sure that rock types have been assigned to all of the units in the model. This will guide you in selecting resistivity values. As discussed previously, the range of resistivities even when the geology is known can be large. Auxiliary information can be used to limit the range. Such information can come from many sources, such as previous electrical or electromagnetic surveys, borehole electric or induction logs, and laboratory electrical measurements of cores or outcrop samples. Auxiliary information in the neighborhood of your study area is preferred; however, it can come from far way. For example, you may be able to find a study from a similar geologic or hydrologic setting on the other side of the world, yet it can provide some guidance in setting resistivity ranges for your model.

We now have a model with features coded as to how certain we are about them, resistivity or ranges of resistivities for the various units in the model, and the features of the model that we want to determine geophysically to answer a specific geologic or hydrologic question. Examples of questions might include things like this: How thick is the unconsolidated overburden? How far from the coast has saltwater intruded into the aquifer? Is the basement rock type granite or sandstone? What is the water quality from the surface to a depth of 50 m? To answer these questions, we will assess the sensitivity of the geophysical response to changes in model parameters (model dimensions and resistivities) associated with the feature we are trying to determine. Changes in model parameters will produce a change in the response, that is, a voltage or corresponding apparent resistivity. If the change in the response is detectable and larger than the expected noise in the data, then the parameter change can be resolved. On the other hand, if the change in the response is insignificant or less than the noise level, the measurement cannot resolve the parameter.

Figure 19 illustrates the model development process. In this example, the situation is a hydrologic investigation needing information on the thickness and lateral extent of a gravelly, sand aquifer. From surficial and bedrock geologic mapping, it is known that the area is covered in most places by a clay unit; however, toward the east, the clay cap is known to be missing because a shale outcrop is shown on a geologic map. The shale dips shallowly to the west. The dipping shale unit is represented as a heavy line in the figure where it crops out, while to the west, it is shown as a dashed line. A water well on the western edge of the study area has a 5 m thick clay unit at the surface, underlain by the gravelly, sand aquifer. At this location, the aquifer extends to a depth of 80 m, at which depth the same shale unit that was mapped to the east is encountered. The interfaces near the well are shown as heavy lines. The clay unit probably exists east of the well based on surficial mapping, so it is shown as a thin, solid line. Because the clay does not extend to the eastern end of the area, the thickness of the clay may decrease to the east or suddenly terminate. If the clay cap is not missing, aquifer recharge can take place through infiltration. Accordingly, this portion of the model shows the base of the clay as a dashed line. The primary question to be answered by the geophysical

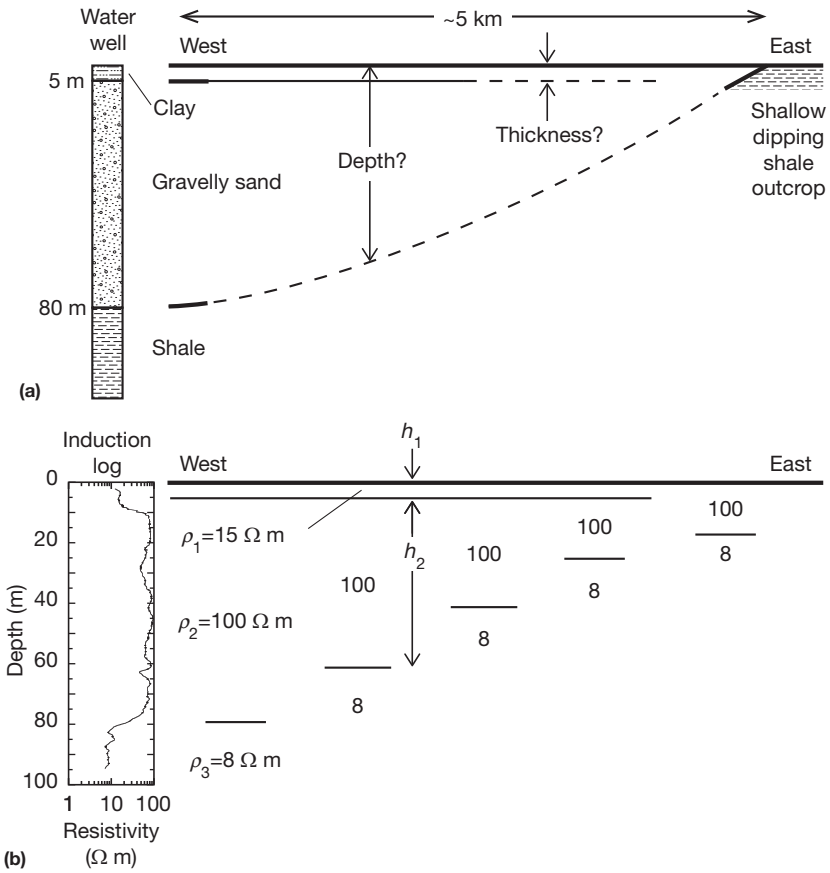


Figure 19 Sketch of exploration model (a) showing known geologic features as heavy solid lines, inferred features as light solid lines, and uncertain features as dashed lines. Known data include the geologically mapped outcrop to the east and lithologic log from a water well to the west. Induction log data (b) from the water well give resistivity estimates for the units. Some of the geophysical models to be tested are shown.

survey is What is the depth to the shale unit?' Secondarily, we want to know the clay thickness.

The second step in the model development process is to assign resistivity values to the units. Because an induction log was measured in the water well, this task is relatively straightforward. The induction log shows three zones: a near-surface conductive zone with a resistivity of 10–20 Ω m, a second zone with a resistivity of 60–100 Ω m, and a final zone at depth with resistivities dropping into the range of 6–12 Ω m. Using these values as guidelines, resistivities of 15, 100, and 8 Ω m were selected for the clay, gravelly sand, and shale, respectively. If a well log did not exist, then resistivity ranges would have to be assigned using the guidelines presented previously in the electrical properties section.

We now have enough information to specify a series of geophysical models corresponding to the sketch (see Figure 19(b)). The electromagnetic response of these models will be calculated using a forward modeling program. Shown is a sequence of three-layer models with the aquifer thickness decreasing to the east and the other model parameters staying fixed. The most easterly model has only two layers because the shale crops out, and the clay unit is presumed absent. Figure 20 shows three scenarios where a different model parameter is varied to see its effect on the apparent resistivity curve. For TEM soundings, measurement noise increases as a function of time after current turnoff. In the earlier parts of the sounding, the apparent resistivity may have 0.1–1% noise, while as time increases, the noise level can increase to over 10%. By the point where the noise level reaches 10%, the data are usually unusable. The width of the horizontal gray line in each plot corresponds to a 10% uncertainty in apparent resistivity. Also shown in each plot is a dashed line that gives the detection limit for the TEM data based on a specified Tx moment (4900 A m^2 for a loop radius of 25 m and a Tx current of 2.5 A) and a noise level of 1 nV m^{-2} (Fitterman, 1989). Data that fall to the left of this line are detectable, while points to the right are not.

The first parameter that we investigate is the thickness of the second layer (h_2) corresponding to the thickness of the gravelly sand in the exploration model (see Figure 20(a)). Five curves are shown with h_2 increasing from 20 to 100 m. In the time range of 1–15 μs, the apparent resistivity decreases rapidly to a minimum value of around 23 Ω m for all curves in response to the 15 Ω m surface layer. At later time, the circulating currents in the ground respond to the resistivity of the second layer, and the apparent resistivity starts to increase. The thicker the 100 Ω m second layer, the larger the increase in apparent resistivity, but in none of the cases shown does it reach the actual layer resistivity. Finally, the apparent resistivity decreases as the current flow is influenced by the 8 Ω m third layer. By the time the curves intersect the detection limit, they have not reached the true third-layer resistivity. The most important point of this graph is the fact that there is a very large separation in the

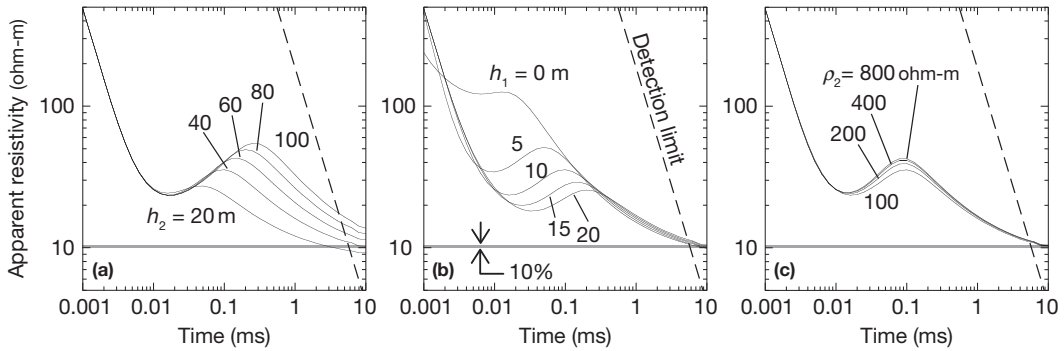


Figure 20 Late-stage apparent resistivity responses for models based on exploration model from **Figure 19**. Transmitter loop radius is 25 m. Assuming a Tx current of 2.5 A and a noise level of 1 nV m^{-2}, the detection limit is shown by the dashed line. The width of the gray horizontal line corresponds to 10% uncertainty in apparent resistivity. The nominal model is $\rho_1 = 15\ \Omega$ m, $h_1 = 10$ m, $\rho_2 = 100\ \Omega$ m, $h_2 = 40$ m, and $\rho_3 = 8\ \Omega$ m. All parameters are nominal values unless otherwise specified. (a) Effect of varying h_2; (b) effect of varying h_1; (c) effect of varying ρ_2.

curves as the second-layer thickness changes, and the difference between each curve is larger than the 10% uncertainty limit. This gives us reason to believe that the TEM sounding can map the depth to the bottom of the aquifer at least in the depth range of 20–100 m.

The second question to be addressed is whether the surface clay layer can be mapped. **Figure 20(b)** shows sounding curves for varying thicknesses of the 8 Ω m first layer. Clearly, there is a detectable difference between the curves, and the absence of the conductive surface layer dramatically changes the curve in the first 0.1 ms. Not all model parameters are resolvable as seen in **Figure 20(c)** where the second-layer resistivity is varied. The nominal model has a ρ_2 value of 100 Ω m. Increasing ρ_2 produces an increase in the apparent resistivity with a maximum near 0.09 ms. The change in apparent resistivity with each doubling of ρ_2 diminishes, and the change between 400 and 800 Ω m is less than the estimated data uncertainty. Thus, we conclude that the presence of the resistive second layer can be established from the TEM soundings; however, if the layer resistivity is too large, we may not be able to estimate its exact value. This shows one of the shortcomings of all EM methods that use inductive sensors – their inability to determine the resistivity of high-resistivity zones. Reliable determination of high-resistivity values requires the use of galvanically coupled methods, such as DC resistivity (see **Chapter 11.08**).

Similar analysis can be used with other EM and geophysical techniques when planning a survey. The importance of model studies cannot be overemphasized as they provide a relatively inexpensive way of assessing the merits and drawbacks of various geophysical techniques. The steps involved are the same; only the techniques and their characteristics are changed. Even if your knowledge of the resistivity values to use in a model is limited and you must use a broad range of values, you will be in a better position to assess the strengths and weaknesses of a particular method and the likelihood of achieving your exploration goal.

11.10.7.1 Sounding Versus Profiling

Electromagnetic geophysical techniques are typically used in one of two modes: *sounding* or *profiling*. The goal of a sounding is to obtain detailed resistivity-depth information at a single location, while the goal of profiling is to map lateral changes in resistivity. Excellent discussions of sounding and profiling can be found in Spies and Frischknecht (1991) and Frischknecht et al. (1991), respectively. A sounding consists of a set of measurements made to obtain information at different depths of exploration. The number of measurements is usually in the tens. In contrast, the number of measurements made at a point on a profile is usually much smaller, from one to less than ten. Consequently, the time required to make a sounding is usually much longer than the time required to make a measurement at a single profile point. Because soundings have a greater number of measurements per location than profiling, the vertical resolution is higher than for profiling. Profiling, on the other hand, provides greater lateral resolution because of the close spacing between measurements. While profiling implies movement from point to point, making soundings is more a stationary endeavor. Accordingly, profiling requires more portable equipment than sounding.

A sounding is made by deploying an array and measuring the response of the ground as a function of some parameter that changes the depth of exploration. The data are plotted as a function of this parameter to give a sense of how the electrical structure of the ground changes with depth. Greater depth of exploration requires that the induced currents flow at a greater depth in the ground. For frequency-domain measurements, the excitation frequency controls the depth of current penetration as expressed by the skin depth δ. If the induced current is flowing deep enough to sense the target zone at depth H ($H/\delta < 1$), changes in frequency have virtually no influence on the depth of exploration, while the intercoil separation does. This is the situation described previously in the section on low-induction number behavior ($B = r/\delta \ll 1$) where a sounding is made by changing the intercoil separation. This technique is called a *geometric sounding*. If the target depth is deeper than the skin depth ($H/\delta > 1$), the induced current does not flow in the target, and the intercoil separation will have little or no effect on the depth of exploration. In this situation, a geometric sounding cannot be made. Changing the frequency, however, does change the depth of exploration, so frequency can be used as the sounding parameter. For time-domain measurements, a similar situation exists when the array size r (intercoil separation or loop radius) is small

compared to the depth to target zone ($r/H < 1$), and time serves as the sounding parameter. Measuring at a later time after current turnoff increases the depth of exploration because the current has diffused deeper into the ground. Changes in array size, on the other hand, have little influence on the depth of exploration other than to increase signal strength.

Profiling employs an array that is moved from point to point along a line. The measurement parameters, be they separation, frequency, or time, are fixed, having been selected beforehand to best detect features in a particular depth range. By having closely spaced measurement points, lateral changes in the subsurface are better resolved. While usually done on the ground, profiling can be performed from an airborne platform. This provides much greater sampling density than ground-based methods and increased speed of data collection. Airborne methods also allow data collection over terrain that is hard to traverse because of vegetation, topography, or land ownership issues.

The selection of sounding or profiling for a particular investigation depends in great part on the goal of the study. Three-dimensional targets are usually more amenable to profiling, whereas one-dimensional targets suggest sounding. The desired interpretative product will also influence the selection. If all that is wanted is a map showing anomalous regions, then profiling may be adequate. This is not to suggest that profile data cannot be interpreted to obtain depth information.

11.10.7.2 Array Types

Another aspect of technique selection is the configuration or array that is used. As we saw for loop–loop EM methods, the coupling to the ground behaves very differently for different orientations of the Tx and Rx coils because the induced current patterns are different. This in turn affects the suitability of the array to a particular target and the depth of exploration. In general, the array should be chosen to induce the largest current possible in the target. For platelike targets, current circulation in the plane of the plate is the goal. For long, thin targets, current flow along the length of the target will produce the largest anomaly. This can be accomplished by inducing current flow in the background material parallel the target's strike direction. These currents are then channeled into the target, assuming that it its more conductive than the surrounding host material. A very good treatment of this extensive subject is given by Frischknecht et al. (1991) for profiling methods and by Spies and Frischknecht (1991) for sounding methods.

11.10.7.3 Noncircular Transmitter Loops

The frequency-domain and time-domain derivations presented earlier in the text have used magnetic dipole transmitters. Only for the time-domain derivation was a nondipolar transmitter loop geometry considered and that was restricted to a circular shape. The obvious advantage of a large transmitter loop is that it provides a larger source moment increasing signal-to-noise ratio in the data and potentially allowing greater depth of exploration. This advantage applies to both frequency-domain and time-domain soundings and is the impetus for the central-induction-loop transient sounding geometry. From a practical standpoint, laying out a circular transmitter loop is not easy, especially at large diameters and when vegetation is present. It is far easier to navigate a straight line than a circle. To this end, we need the ability to compute the frequency-domain and time-domain responses of large square (or rectangular) loops. The simplest solution to this problem is to replace the square loop with a circular loop of the same area. This works well at later times; however, it will introduce some errors at earlier times. A better approach for polygonal loops is to divide the loop into an equivalent set of elemental electric dipoles. The magnetic field produced by each electric dipole at the receiver coil is computed after applying the appropriate reciprocity transformation and by integrating the contribution of all of the dipoles forming the transmitter loop. Ryu et al. (1970) developed this approach for a circular loop and Poddar (1982, 1983a,b) for a rectangular loop in the frequency domain. Raiche (1987) provided an analysis of the problem for any polygonal transmitter loop for both the frequency and time domains. Raiche's formulation is unique in that it uses a numerical inverse Laplace transform to compute the time-domain response. Alternatively, Ward and Hohmann (1988) used a Fourier transform representation of the magnetic dipole source in the frequency domain and integrated it over the area of the transmitter loop. The large loop, time-domain response then is computed by taking a cosine (or sine) Fourier transform of the frequency-domain response as described previously.

11.10.8 Applications and Examples

In this section, a number of examples on the use of EM methods for near-surface investigations are presented. The examples presented are not meant to be exhaustive, but rather to give a sampling of some typical and unusual applications.

11.10.8.1 Saltwater Intrusion Mapping

In many parts of the world, groundwater is the principle source of drinking water. Groundwater supply wells are subject to contamination from man-made and natural sources. In coastal areas relying on groundwater, saltwater intrusion is of great concern because once seawater gets into a formation, it is nearly impossible to flush it out. Over pumping of supply wells is a common cause of saltwater intrusion. As water is pumped from a well, water in the aquifer flows toward the well. If flow paths exist between the aquifer and the ocean, seawater can be drawn toward the well. Prudent management of coastal, groundwater resources requires some sort of monitoring program to ensure that drinking water supplies are not in jeopardy. Monitoring can be accomplished by routinely measuring the water quality in the supply well; however, this may not give adequate warning to change pumping rates before irreversible contamination occurs. The installation of dedicated monitoring wells between supply wells and the ocean greatly improves the situation. Monitoring can be done by pumping samples from screened intervals or by using induction logs run in the well. The advantage of induction logs is that they can be made throughout the entire length of the well, not just the screened interval, and the induction tool senses beyond the wellbore (Barlow, 2003; Hittle, 1999). Monitoring wells are not without their own problems; that is, they are expensive and stationary. Electromagnetic measurements are

the perfect adjunct to monitoring wells because they are much less expensive, and they can be deployed over a region to accurately locate the saltwater front and monitor its movement.

We will look at an example of the use of ground-based TEM soundings and a HEM survey to map saltwater intrusion in the Biscayne aquifer located in Miami-Dade County, Florida (Fitterman and Prinos, 2011; Fitterman et al., 2012). The Biscayne aquifer is the upper aquifer in the surficial aquifer system, and the gray limestone aquifer is the lower aquifer. The base of the surficial aquifer system is the intermediate confining unit of the Hawthorn Group (Fish and Stewart, 1991). The Biscayne aquifer is a shallow, highly transmissive, limestone aquifer found in eastern Miami-Dade County that is virtually clay-free (Fish and Stewart, 1991). The relationships between chloride content, water resistivity, and formation resistivity for the Biscayne aquifer were shown previously in **Figures 1** and **2**. Formation resistivities of less than 10 Ω m are considered to have a chloride content of more than 1000 mg l^{-1} indicating contamination with seawater. The freshwater-saturated Biscayne aquifer will have resistivities of greater than 30 Ω m.

TEM soundings were used throughout Miami-Dade County in the area shown in **Figure 21**. A line of four TEM soundings that cross the saltwater front is shown in **Figure 22**. The location of the soundings relative to the saltwater front can be seen in the strip map in the **Figure 23(d)**. Soundings MIA209 and MIA210 are located on the freshwater side of the front. Their apparent resistivity-time curves show a monotonic decrease with increasing time to about 0.4 ms. The inverted resistivity-depth functions show a high-resistivity (>100 Ω m) surface layer, which is underlain by a moderately resistive zone highlighted in blue. The second layer corresponds to the freshwater-saturated Biscayne aquifer. The time range of the TEM sounding corresponding approximately to this layer is also highlighted in blue. Soundings MIA211 and MIA212, which are on the saltwater-intruded side of the interface, have a well-defined minimum marked in red. In the resistivity-depth model, the minimum corresponds to the low-resistivity zone (<10 Ω m), which is caused by saltwater in the Biscayne aquifer. Below the saltwater-intruded zone, there is a higher resistivity, freshwater zone in the gray limestone aquifer, which underlies the Biscayne aquifer. TEM soundings similar to these were used throughout Miami-Dade County to refine the location of the saltwater front.

A HEM survey was flown in the Model Land Area southeast of Florida City (see **Figure 21**) to map saltwater intrusion. Flight lines spaced 400 m apart were flown in the east–west

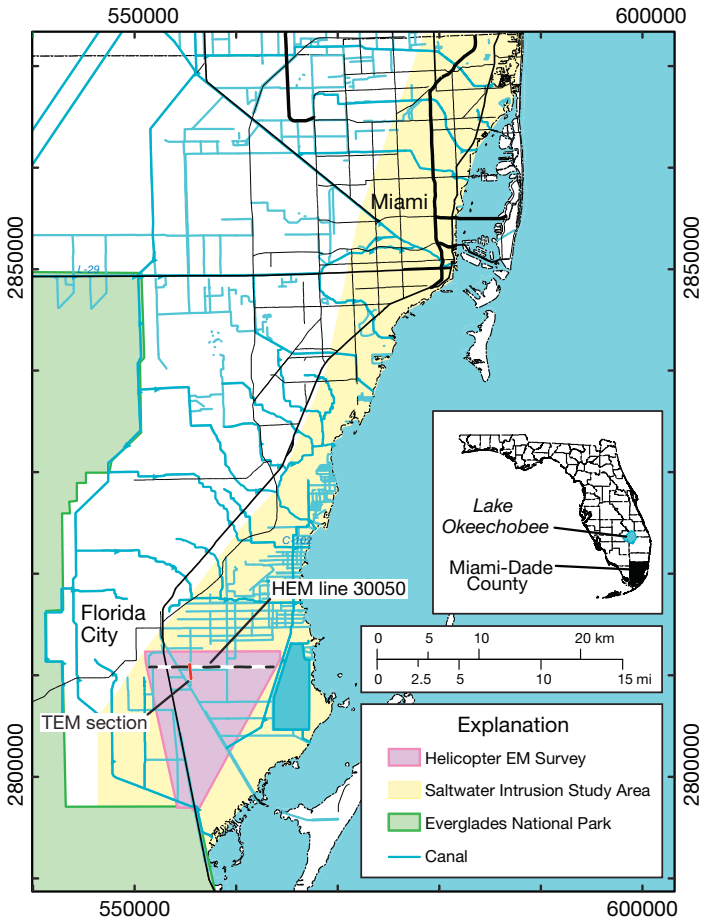


Figure 21 Location of saltwater intrusion study and helicopter electromagnetic survey in south Florida. The TEM section is shown in **Figure 22**, and HEM line 30050 is shown in **Figure 23**. Modified from Fitterman DV, Deszcz-Pan M, and Prinos ST (2012) Helicopter electromagnetic survey of the Model Land area, Southeastern Miami-Dade County, Florida. U.S. Geological Survey. *Open-File Report 2012-1176*.

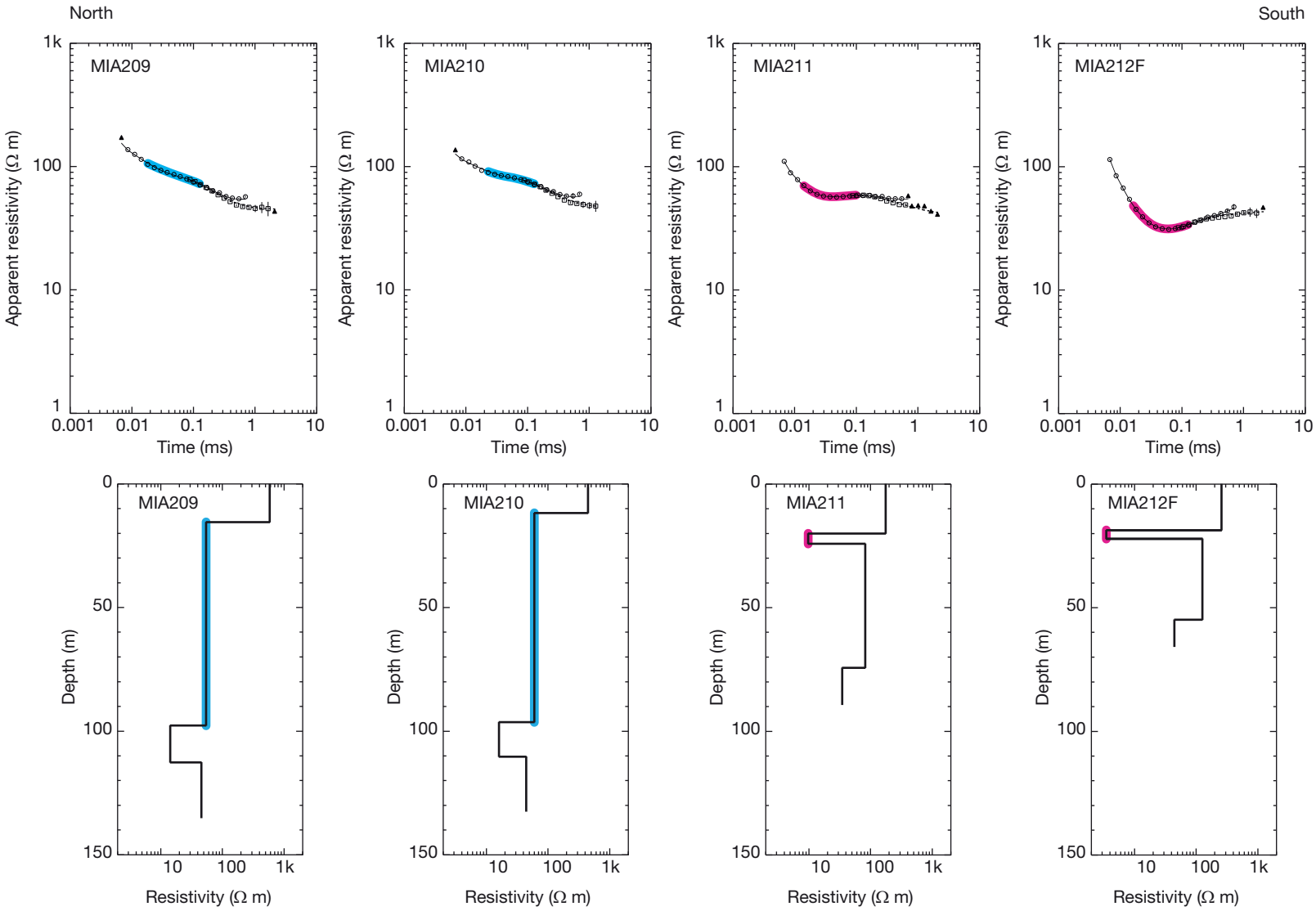


Figure 22 Time-domain EM sounding curves (upper row) and corresponding interpreted resistivity-depth models (lower row). Locations of the soundings in south Florida are shown in **Figures 21** and **23(d)**. The saltwater front is located between sounding MIA210 and MIA211. Blue and red highlights in resistivity-depth model show the freshwater and saltwater zones, respectively. Corresponding features in the sounding curves are similarly highlighted. Modified from Fitterman DV and Prinos ST (2011) Results of time-domain electromagnetic soundings in Miami-Dade and southern Broward Counties, Florida. U.S. Geological Survey. *Open-File Report 2011-1299*.

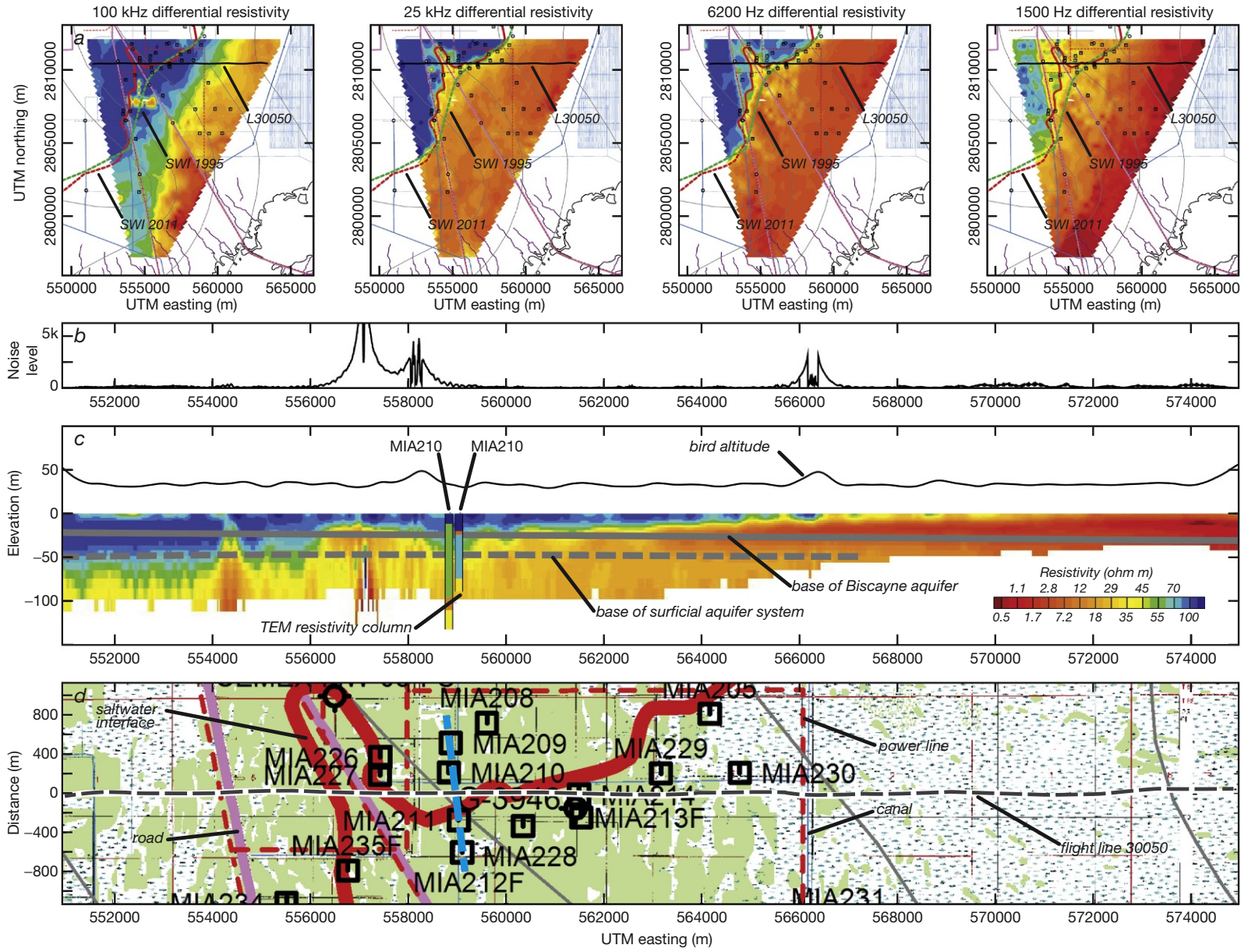


Figure 23 Helicopter EM data from Miami-Dade County, Florida. Panel (a) shows the apparent resistivity maps for four measured frequencies. (The 400 Hz apparent resistivity map is not shown.) The saltwater interface (SWI) as determined in 1995 (Sonenshein, 1997) and 2011 (Fitterman et al., 2012) is shown. Panel (b) shows the uncalibrated EM noise due to power lines. Panel (c) shows the inverted resistivity cross section. Nearby TEM interpretations are shown as colored columns. Resistivity color scale is applicable to apparent resistivity maps, cross section, and TEM interpretations. Panel (d) shows HEM flight line, saltwater interface, and TEM locations (black squares) superimposed on topographic map. The line of soundings in **Figure 22** is marked by a dashed blue line. Modified from Fitterman DV, Deszcz-Pan M, and Prinos ST (2012) Helicopter electromagnetic survey of the Model Land area, Southeastern Miami-Dade County, Florida. U.S. Geological Survey. *Open-File Report 2012-1176*.

direction approximately perpendicular to the coast. Along flight lines, measurement spacing was about 3 m. Flight-line spacing influences the detail in the resulting map products. The lateral resolution or footprint of the measurements depends on several factors (Beamish, 2003). The footprint varies linearly with the elevation of the transmitter coil and to first order is about twice the bird elevation. Horizontal Tx loops have a larger footprint by a factor of 1.3–1.5 than vertical Tx loops. The measurement frequency and the resistivity of the ground affect the footprint to a smaller degree. The bird contained five HCP coil pairs operating at frequencies between 400 Hz and 100 kHz. With a nominal bird elevation of 30 m, the footprint is a square of about 100–140 m on a side (Beamish, 2003). The raw data are converted to differential apparent resistivity at each frequency. Maps for four frequencies are shown in **Figure 23(a)**. The depth of exploration increases as the frequency decreases. The 100 kHz map shows saltwater at shallow depths as a 5 km wide, orange and yellow band subparallel the coastline. Two low-resistivity features (orange patches in the blue, high-resistivity zone) are associated with two quarries. The formation factor in the quarry pit is 1 because the rock has been removed, and the bulk resistivity, that of the water alone, is lower than in the surrounding, unmined areas. As the frequency decreases, the depth of exploration increases, and the 10 Ω m transition moves landward. In the 25 kHz and 6200 Hz maps, the measurements sense to the base of the Biscayne aquifer, indicating saltwater saturation at its base. In both maps, a prong of saltwater aligned with the Card Sound Road Canal and Card Sound Road oriented at an azimuth of 330° is seen. Card Sound Road was built in the 1920s using material excavated from the adjacent canal. Tidal flow in the canal was uncontrolled as recently as 2012 allowing inland saltwater flow as far north as Florida City during storm surges and periods of drought when the water table is depressed.

Power lines in the survey area cause two problems. First, the bird must be flown high enough so that it does not snag the power line, and second, the power line is a noise source that degrades data quality. **Figure 23(b)** shows the power-line noise amplitude recorded in the bird. The peaks in the noise trace correspond to locations where the flight line crosses three power lines, which are shown as dashed red lines in **Figure 23(d)**. The power lines near UTM E 558000 and E 566000 are high enough that the helicopter must increase altitude as seen in the bird altitude trace (**Figure 23(c)**). The HEM data are inverted as 1-D models that are stitched together to produce the interpreted resistivity cross section (**Figure 23(c)**). On the eastern end of the section, the resistivities are very low due to saltwater intrusion through the entire depth of the Biscayne aquifer. Saltwater is detected at the base of the Biscayne as far west as UTM E 560200. Between UTM E 556300 and UTM E 558300, the low-resistivity zone corresponds to the saltwater tongue associated with the Card Sound Road Canal. Noise from the power line near UTM E 557000 produces vertical striping in the inversion results. The effect was not seen near other power lines where the noise level was less. Near UTM E 554200, there is a rise in the low resistivity from the bottom of the surficial aquifer to just below the bottom of the Biscayne. This anomaly is produced by relict seawater that flowed up ditches on both sides of the former train grade that occupied the present-day location of US Highway 1. The train grade was built in 1905, and the uncontrolled ditches allowed landward seawater flow during storms and droughts. By 1944, the ditches were blocked. Seawater that had previously flowed inland sank to the bottom of the aquifer over time and remains there to this day producing a low-resistivity zone.

The saltwater interface shown in **Figure 23(d)** was drawn using the results from the interpreted HEM cross section augmented with the TEM data. In general, the correspondence between the HEM and TEM results is very good. Interpreted TEM resistivity-depth columns for soundings MIA210 and MIA211 are superimposed on the HEM section. MIA210 is landward of the saltwater front and does not show the low-resistivity zone seen in MIA211, which is seaward of the saltwater front. The locations of the saltwater interface in 1995 and 2011 are shown on the apparent resistivity maps.

11.10.8.2 Coastal Bathymetry (AEM)

Airborne electromagnetic bathymetry provides a rapid alternative to shipborne measurement of water depths in coastal areas. Since its suggestion by Morrison and Becker (1982), extensive research has refined the method. Initial bathymetric AEM surveys used the INPUT time-domain system flown on a fixed-wing platform (Zollinger et al., 1987). The depth of exploration as great as 40 m with 2 m resolution was demonstrated. HEM systems were also employed with 1 m resolution in water depths of 13–20 m (Won and Smits, 1987, 1990). More recently, a helicopter-borne, time-domain system specifically designed for bathymetry has been developed (Vrbancich et al., 2010). Employed in Sydney Harbour (NSW, Australia), it has mapped water depth, bottom sediment thickness, and depth to bedrock in an area with complex geology produced by sea-rise flooding of sediments deposited in coastal paleovalleys (Vrbancich, 2012). Using auxiliary data in the form of water column conductivity, bathymetric accuracies of 5% to 10% to depths of 30 m were achieved using fixed layer resistivities in the data inversion.

11.10.8.3 Sea-Ice Thickness Mapping

Another ocean application is the mapping of sea-ice thickness using HEM (Becker et al., 1983; Kovacs and Valleau, 1990). The resulting map can be used for navigation through pack ice along the path of thinnest ice and for global climate change studies. The resistivity of sea ice is relatively high, making it nearly transparent to EM energy, while the conductive seawater on which it floats is a very good EM target. The EM data are used to estimate the distance between the EM bird and the seawater, and a laser altimeter is used to determine the height of the bird above the ice. Assuming that the ice forms as planar sheets, the difference between these distances gives the ice thickness. In reality, this assumption is rarely true. Lateral stresses in the ice caused by wind and wave action produce pressure ridges, which are underlain by ice keels. These features can be very three-dimensional, complicating data interpretation. Pfaffhuber et al. (2012) reported on a HEM bird specifically designed for sea-ice mapping incorporating multiple Tx–Rx coil pairs, LIDAR ice topography measurements, and GPS receivers. While still in development, this bird holds promise for accurate 3-D data interpretation. Reid et al.

(2003) described the use of a terrain conductivity meter (EM-31) mounted on an 11 m boom outboard of an icebreaker. The height of the EM-31 above the ice is measured with a laser altimeter. Seawater conductivity is also measured independently so that the sea-ice thickness is the only unknown. With the EM-31 2–5 m above the surface, data inversion can determine ice thicknesses of up to 2.5 m.

11.10.8.4 Clay Mapping

Clay deposits play an important role in hydrologic models due to their low permeability. Hydrologic models are critically dependent upon knowing the location, lateral extent, and thickness of clay zones. Obtaining this subsurface information from wells alone is often not practical because the subsurface geology can change between well sites as discussed before in the saltwater intrusion example from south Florida. Fortunately, clay deposits are conductive making them ideal EM targets. This example illustrates the use of TEM soundings to map the blue clay found in the San Luis Valley (SLV) of south central Colorado (see **Figure 24**). The blue clay separates an unconfined near-surface aquifer from a deeper, confined, artesian aquifer. The blue clay was deposited at the bottom of Lake Alamosa that formed in the SLV about 3–3.5 Ma (Machette et al., 2007). Lake Alamosa overflowed and incised the volcanic rocks at the southern end of the valley around 440 ka. The location of the blue clay is an important feature of the regional hydrologic model, and its location must be known for permitting of agricultural and domestic wells in the SLV.

The SLV is the home of Great Sand Dunes National Park and Monument and Preserve (GRSA), which is located along the eastern edge of the valley. Early in this century, commercial interests sought to purchase groundwater rights in the SLV to export water to municipalities in other parts of Colorado. The National Park Service (NPS), which administers the GRSA, became concerned that proposed supply wells near the GRSA would endanger groundwater levels in the park and possibly the sand dunes. To protect the park's interests, the NPS sought and received a nonuse water right that prevented wells outside the GRSA from lowering the water table inside the GRSA (Colorado District Court Water Division No. 3, 2008). As part of the court decree, the NPS was required to install a system of piezometer wells along the western boundary of the GRSA. The wells needed to extend to the blue clay, or other adequately thick clay zones, so that the piezometers accurately reflected the condition of the unconfined aquifer.

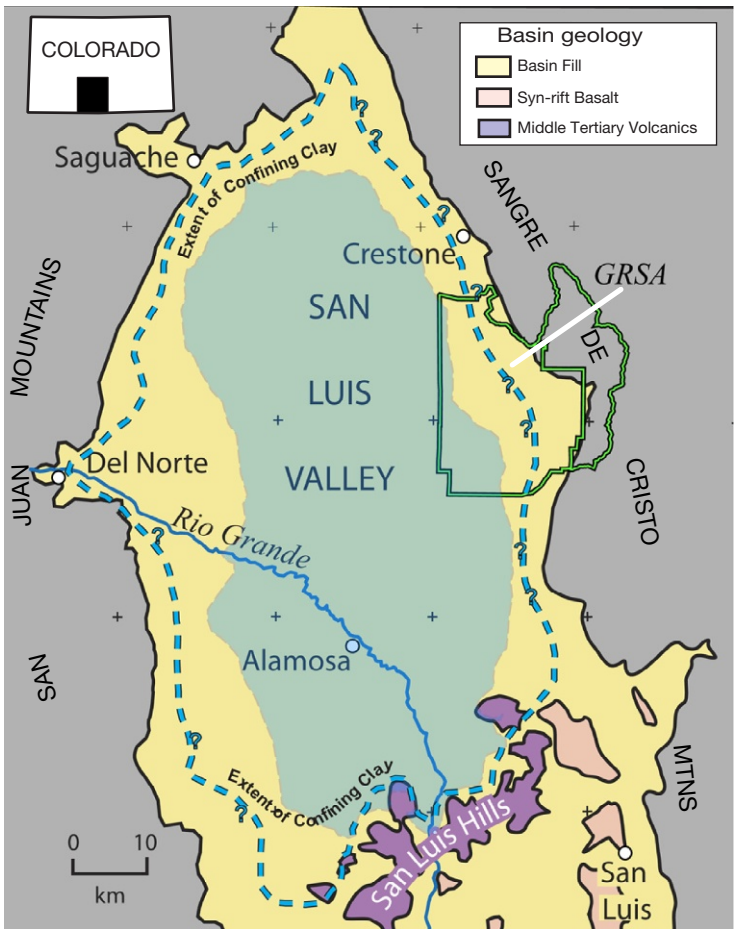


Figure 24 Map of San Luis Valley (shaded green) in southern Colorado showing extent of blue clay as dashed line (Machette et al., 2007). Location of Great Sand Dunes National Monument and Preserve (GRSA) is shown.

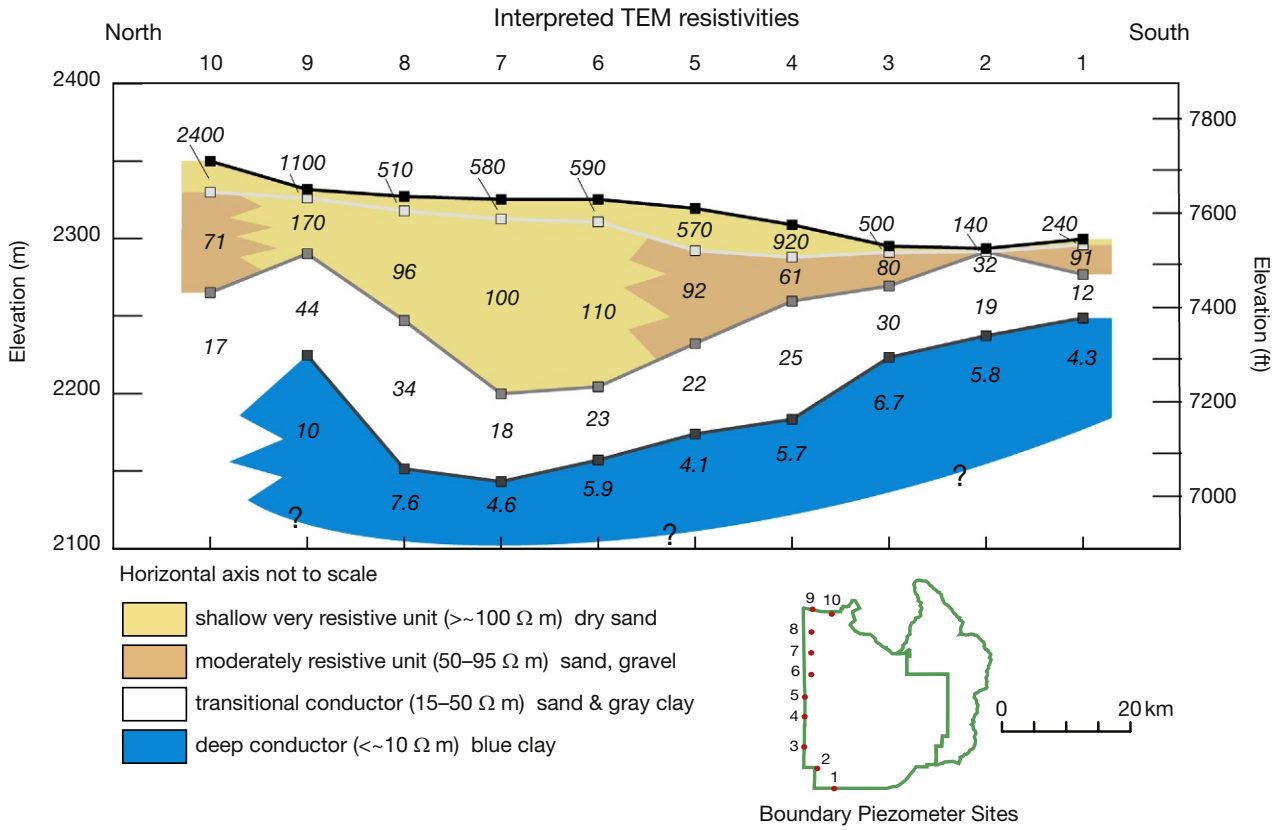


Figure 25 Interpreted time-domain EM soundings along the western boundary of Great Sand Dunes National Monument and Preserve (GRSA), Colorado. TEM locations are displayed as evenly spaced in pseudo-cross section, though actual spacing varies. Horizontal scale applies only to index map.

Ten piezometer sites were selected along the western park boundary (see Figure 25). To plan the drilling program, TEM soundings were made at the piezometer sites by the U.S. Geological Survey to determine the depth to clay.

Analysis of lithologic and induction logs from previously drilled wells near the GRSA was used to determine the resistivity of materials found in the unconfined aquifer including sand, gray clay, and blue clay (Fitterman and Grauch, 2010). The average sand resistivity is 177 Ω m with values ranging from 32 Ω m to over 1 k Ω m. Gray clay resistivity averages around 21.4 Ω m with values ranging from 8 to 64 Ω m. The blue clay is significantly more conductive with an average resistivity of 6.9 Ω m with a range of 4–12 Ω m. The large difference in resistivity of these materials, in particular the factor of three differences in resistivity between the gray and blue clays, means that zones composed of these materials can be discriminated by geophysical measurements.

A pseudo-cross section of the inverted TEM soundings from the piezometer sites is shown in Figure 25. (The cross section is termed *pseudo* because the interpretations are displayed as evenly spaced, whereas the actual spacing varies slightly and the soundings are not in a straight line.) Each interpretation is represented as a vertical column with the layer resistivities given. Resistivity values are color-coded using ranges that separate the results into hydrologically significant groupings. The general pattern is a very resistive dry sand (>100 Ω m) at the surface underlain by moderately resistive sand and gravel (50–95 Ω m). These units are underlain by a transitional conductor in the 15–50 Ω m range composed of sand and gray clays. A very conductive layer (<10 Ω m) is found at depth at all but one location. This unit is interpreted as the blue clay. Moving across the section from north to south, the surface elevation decreases and the blue clay becomes shallower. These and other TEM soundings made throughout the GRSA and on adjacent land reveal that there is a great deal of relief on the top of the blue clay. This can be due to either pre-lake era topography or post-depositional deformation. The exact cause is still under investigation. Comparison of the drilling results with the TEM interpretations gave excellent agreement.

11.10.8.5 Airborne Mapping of UXO

Mapping of UXO has become necessary as military facilities are closed and the land is put to other uses. Ground-based and airborne magnetic methods provide a rapid and low-cost screening technique. In volcanic terrains, the level of background magnetic noise can be so large as to render magnetic measurements ineffective (Doll et al., 2012). EM methods provide an alternative method that can be used separately or in conjunction with magnetic methods. Doll et al. (2010) described a time-domain EM system composed of two

transmitters and eight receivers in a 3 m by 12 m, nonconducting structure that is mounted on the skids below a helicopter. The Tx–Rx array is flown at a height of only 1.5–3 m above the ground to ensure high target sensitivity. Field tests have shown this system capable of mapping 81 mm mortar shells in basaltic areas with moderate to severe geologic noise, while magnetic measurements were unreliable.

11.10.9 Summary

The use of EM techniques has evolved since they were first developed for mineral exploration in the 1930s. Success with early systems fostered improvement of instrumentation and interpretation. Modern electronics allowed for more compact and reliable instrumentation. As computers became less expensive and faster, numerical interpretation methods have improved. These advances made EM geophysics a key component of mineral exploration. Concomitant with the maturation of EM methods, users started to think of other applications. Groundwater investigations were one of the first nonmineral areas to see the use of EM methods. Their use, which was rare at first, has become commonplace. The progression from groundwater to environmental applications was a natural transition. Many environmental targets are much shallower than mineral and groundwater ones, requiring increased sensitivity and resolution. Hardware developers have addressed this need with equipment intended for targets within 1–2 m of the surface. This constant push and pull by the users and developers of the techniques has led to very sophisticated and reliable EM geophysical methods with ever-expanding applications.

References

Abu Rajab J and El-Naqa A (2013) Mapping groundwater salinization using transient electromagnetic and direct current resistivity methods in Azraq Basin, Jordan. *Geophysics* 78: B89–B101.

Anderson WL (1975) Improved digital filters for evaluating Fourier and Hankel transform integrals. U.S. Geological Survey. *Open-File Report 75-12*.

Anderson WL (1979) Numerical integration of related Hankel transforms of orders 0 and 1 by adaptive digital filtering. *Geophysics* 44: 1287–1305.

Angoran YE, Fitterman DV, and Marshall DJ (1974) Induced polarization: A geophysical method for locating cultural metallic refuse. *Science* 184: 1287–1288.

Annan AP (1974) *The Equivalent Source Method for Electromagnetic Scattering Analysis and Its Geophysical Application*. PhD Thesis, Memorial University of Newfoundland.

Archie GE (1942) The electrical resistivity log as an aid to determining some reservoir characteristics. *Transactions of the American Institute of Mining, Metallurgical and Petroleum Engineers* 146: 54–62.

Asch TH, Wright DL, Moulton CW, Irons TP, and Nabighian MN (2008) ALLTEM UXO detection and discrimination. *SEG Technical Program Expanded Abstracts 2008*.

Auken E and Christiansen A (2004) Layered and laterally constrained 2D inversion of resistivity data. *Geophysics* 69: 752–761.

Auken E, Christiansen AV, Jacobsen L, and Sørensen KI (2005) Laterally constrained 1D-inversion of 3D TEM data. In: *Symposium on the Application of Geophysics to Engineering and Environmental Problems 2005*. Tulsa: Society of Exploration Geophysicists.

Barlow PM (2003) Ground water in fresh water-salt water environments of the Atlantic Coast. U.S. Geological Survey. U.S. Geological Survey Circular 1262.

Barringer AR (1990) Historical aspects of airborne electromagnetic development. In: Fitterman DV (ed.) *Developments and Applications of Modern Airborne Electromagnetic Surveys*, pp. 7–8. U.S. Geological Survey, U.S. Geological Survey Bulletin 1925.

Beamish D (2003) Airborne EM footprints. *Geophysical Prospecting* 51: 49–60.

Becker A, Barringer AR, and Annan AP (1990) Airborne electromagnetics 1978–1988. In: Fitterman DV (ed.) *Developments and Applications of Modern Airborne Electromagnetic Surveys*. U.S. Geological Survey, U.S. Geological Survey Bulletin 1925.

Becker A, Morrison H, and Smits K (1983) Analysis of airborne electromagnetic systems for mapping thickness of sea ice. *Naval Ocean Research and Development Activity NORDA Technical Note 216*.

Brace WF and Orange AS (1968) Further studies of the effects of pressure on electrical resistivity of rocks. *Journal of Geophysical Research* 73: 5407–5420.

Chave AD, Constable SC, and Edwards RN (1991) Electrical exploration methods for the seafloor. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics—Applications*, vol. 2B. Tulsa: Society of Exploration Geophysicists.

Christiansen A and Christensen N (2003) A quantitative appraisal of airborne and ground-based transient electromagnetic (TEM) measurements in Denmark. *Geophysics* 68: 523–534.

Colorado District Court Water Division No. 3 (2008) In the matter of the application for water rights of the United States of American in Alamosa and Saguache Counties. Case Number: 2004CW35.

Constable S (2006) Marine electromagnetic methods—A new tool for offshore exploration. *The Leading Edge* 25: 438–444.

Constable S and Srnka L (2007) An introduction to marine controlled-source electromagnetic methods for hydrocarbon exploration. *Geophysics* 72: WA3–WA12.

Dickinson J, Pool D, Groom R, and Davis L (2010) Inference of lithologic distributions in an alluvial aquifer using airborne transient electromagnetic surveys. *Geophysics* 75: WA149–WA161.

Doll WE, Gamey TJ, Bell DT, et al. (2012) Historical development and performance of airborne magnetic and electromagnetic systems for mapping and detection of unexploded ordnance. *Journal of Environmental & Engineering Geophysics* 17: 1–17.

Doll W, Gamey T, Holladay J, et al. (2010) Results of a high-resolution airborne TEM system demonstration for unexploded ordnance detection. *Geophysics* 75: B211–B220.

Draskovits P and Verö L (2005) Indication of groundwater contamination with the induced-polarization (IP) method. In: Butler DK (ed.) *Near Surface Geophysics, Part 2: Applications and Case Histories*. Tulsa: Society of Exploration Geophysicists.

Ellefsen KJ, Lucius JE, and Fitterman DV (1998) An evaluation of several geophysical methods for characterizing sand and gravel deposits. U.S. Geological Survey. *Open-File Report 98-221*.

Fish JE and Stewart M (1991) Hydrogeology of the surficial aquifer system, Dade County, Florida. U.S. Geological Survey. *Water-Resources Investigation Report 90-4108*.

Fitterman DV (1986) Transient electromagnetic sounding in the Michigan Basin for ground water evaluation. In: *NWWA Conference on Surface and Borehole Geophysical Methods and Ground Water Instrumentation, October 15–17, Denver, Colorado*, pp. 334–353. Dublin, OH: National Water Well Association.

Fitterman DV (1987) Examples of transient sounding for ground-water exploration in sedimentary aquifers. *Ground Water* 25: 685–692.

Fitterman DV (1989) Detectability levels for central induction transient soundings. *Geophysics* 54: 127–129.

Fitterman DV, Menges CM, Al Kamali AM, and Jama FE (1991) Electromagnetic mapping of buried paleochannels in eastern Abu Dhabi Emirate, U.A.E. *Geoexploration* 27: 111–133.

Fitterman DV, Deszcz-Pan M, and Stoddard CE (1999) Results of time-domain electromagnetic soundings in Everglades National Park, Florida (on CD-ROM). U.S. Geological Survey. *Open-File Report 99-426*.

Fitterman DV, Deszcz-Pan M, and Prinos ST (2012) Helicopter electromagnetic survey of the Model Land area, Southeastern Miami-Dade County, Florida. U.S. Geological Survey. *Open-File Report 2012-1176*.

Fitterman DV (2014) Mapping saltwater intrusion in the Biscayne aquifer, Miami-Dade County, Florida using transient electromagnetic sounding. *Journal of Environmental and Engineering Geophysics* 19(1): 33–43.

Fitterman DV and Grauch VJS (2010) Transient electromagnetic mapping of clay units in the San Luis Valley, Colorado. In: *Proceedings of the 23rd Symposium on the Application of Geophysics to Engineering and Environmental Problems SAGEEP 2010, April 11–15, 2010*, pp. 154–164. Keystone, Colorado: Environmental and Engineering Geophysical Society.

Fitterman DV and Labson VF (2005) Electromagnetic induction methods for environmental problems. In: Butler DK (ed.) *Near Surface Geophysics, Part 1: Concepts and Fundamentals*, pp. 301–356. Tulsa: Society of Exploration Geophysicists.

Fitterman DV and Prinos ST (2011) Results of time-domain electromagnetic soundings in Miami-Dade and southern Broward Counties, Florida. U.S. Geological Survey. *Open-File Report 2011–1299*.

Fraser DC (1972) A new multicoil aerial electromagnetic prospecting system. *Geophysics* 37: 518–537.

Fraser DC (1978) Resistivity mapping with an airborne multicoil electromagnetic system. *Geophysics* 43: 144–172.

Fraser DC (1979) The multicoil II airborne electromagnetic system. *Geophysics* 44: 1367–1394.

Freeze RA and Cherry JA (1979) *Groundwater*. Englewood Cliffs, NJ: Prentice Hall.

Frischknecht FC (1967) Fields about an oscillating magnetic dipole over a two-layer earth. *Quarterly of the Colorado School of Mines* 62: 326.

Frischknecht FC (1988) Electromagnetic physical scale modeling. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1. Tulsa: Society of Exploration Geophysicists.

Frischknecht FC, Labson VF, Spies BR, and Anderson WL (1991) Profiling methods using small sources. In: Nabighian MN (ed.). *Electromagnetic Methods in Applied Geophysics—Applications*, vol. 2A. Tulsa: Society of Exploration Geophysicists.

Goldman M, Hurwitz S, Gvirtzman H, Rabinovich B, and Rotstein Y (1996) Application of marine time-domain electromagnetic method in lakes: The Sea of Galilee, Israel. *European Journal of Environmental and Engineering Geophysics* 1: 125–138.

Goldstein MA and Strangway DW (1975) Audio-frequency magnetotellurics with a grounded electric dipole source. *Geophysics* 40: 669–683.

Grant FS and West GF (1965) *Interpretation Theory in Applied Geophysics*. New York: McGraw-Hill.

Griffiths DJ (1999) *Introduction to Electrodynamics*. Upper Saddle River, NJ: Prentice Hall.

Grimm RE, Olhoeft GR, McKinley K, Rossabi J, and Riha B (2005) Nonlinear complex-resistivity survey for DNAPL at the Savannah River Site A-014 Outfall. *Journal of Environmental and Engineering Geophysics* 10: 351–364.

Hanna WF (1990) Some historical notes on early magnetic surveying in the U.S. Geological Survey. In: Hanna WF (ed.) *Geologic Applications of Modern Aeromagnetic Surveys*: U.S. Geological Survey, U.S. Geological Survey Bulletin 1924.

Hatch M, Munday T, and Heinson G (2010) A comparative study of in-river geophysical techniques to define variations in riverbed salt load and aid managing river salinization. *Geophysics* 75: WA135–WA147.

Hearst JR and Nelson PH (1985) *Well Logging for Physical Properties*. New York: McGraw-Hill.

Hem JD (1970) Study and interpretation of the chemical characteristics of natural water. U.S. Geological Survey. Water-Supply Paper 1473.

Hittle CD (1999) Delineation of saltwater intrusion in the surficial aquifer system in eastern Palm Beach, Martin, and St. Lucie Counties, Florida, 1997–98. U.S. Geological Survey. *Water-Resources Investigations Report 99-4214*.

Hohmann GW (1988) Numerical modeling for electromagnetic methods of geophysics. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics—Applications*, vol. 1. Tulsa: Society of Exploration Geophysicists.

Hohmann GW and Raiche AP (1988) Inversion of controlled-source electromagnetic data. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1. Tulsa: Society of Exploration Geophysicists.

Huang H and Fraser DC (1996) The differential parameter method for multifrequency airborne resistivity mapping. *Geophysics* 61: 100–109.

Jackson JD (1962) *Classical Electrodynamics*. New York: John Wiley & Sons, Inc.

Jackson DB and Keller GV (1972) An electromagnetic sounding survey of the summit of the Kilauea Volcano, Hawaii. *Journal of Geophysical Research* 66: 4957–4965.

Jackson PD, Smith DT, and Stanford PN (1978) Resistivity–porosity–particle shape relationships for marine sands. *Geophysics* 43: 1250–1268.

Johansen HK (1977) A man/computer interpretation system for resistivity soundings over a horizontally stratified earth. *Geophysical Prospecting* 25: 667–691.

Johansen HK and Sorensen K (1979) Fast Hankel transforms. *Geophysical Prospecting* 27: 876–901.

Kaufman A (1978) Frequency and transient responses of electromagnetic fields created by currents in confined conductors. *Geophysics* 43: 1002–1010.

Kaufman AA and Keller GV (1983) *Frequency and Transient Sounding*. Amsterdam: Elsevier.

Keller GV (1971) Natural-field and controlled-source methods in electromagnetic exploration. *Geoexploration* 9: 99–147.

Keller GV (1988) Rock and mineral properties. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1. Tulsa: Society of Exploration Geophysicists.

Keller GV and Frischknecht FC (1966) *Electrical Methods in Geophysical Prospecting*. Oxford: Pergamon.

Knight RJ and Endres AL (2005) An introduction to rock physics principles for near-surface geophysics. In: Butler DK (ed.) *Near Surface Geophysics, Part 1: Concepts and Fundamentals*. Tulsa: Society of Exploration Geophysicists.

Kovacs A and Valleau NC (1990) Airborne electromagnetic measurement of sea ice thickness and sub-ice bathymetry. In: Fitterman DV (ed.) *Developments and Applications of Modern Airborne Electromagnetic Surveys*: U.S. Geological Survey, U.S. Geological Survey Bulletin 1925.

Lamontagne Y and West G (1971) EM response of a rectangular thin plate. *Geophysics* 36: 1204–1222.

Machette MN, Marchetti DW, Thompson RA (2007) Ancient Lake Alamosa and the Pliocene to middle Pleistocene evolution of the Rio Grande (Chapter G). In: Machette MN, Coates M-M, Johnson ML (eds.) *Rocky Mountain Section Friends of the Pleistocene Field Trip—Quaternary Geology of the San Luis Basin of Colorado and New Mexico*, September 7–9, 2007. U. S. Geological Survey. *Open-File Report 2007-1193*.

Macnae J, Lamontagne Y, and West G (1984) Noise processing techniques for time-domain EM systems. *Geophysics* 49: 934–948.

Madden TR and Marshall DJ (1967) Part D—Induced polarization, a review. In: Hansen DA, Heinrichs WE Jr, Holmer RC, Macdougall RE, Rogers GR, Sumner JS, and Ward SH (eds.) *Mining Geophysics*. Tulsa: Society of Exploration Geophysicists.

March H (1953) The field of a magnetic dipole in the presence of a conducting sphere. *Geophysics* 18: 671–684.

Marshall DJ and Madden TR (1959) Induced polarization, a study of its causes. *Geophysics* 24: 790–816.

Maxwell JC (1873) *A Treatise on Electricity and Magnetism*. Oxford: Clarendon Press.

McCracken K, Oristaglio M, and Hohmann G (1986) Minimization of noise in electromagnetic exploration systems. *Geophysics* 51: 819–832.

Mcneill JD (1990) Use of electromagnetic methods for groundwater studies. In: Ward SH (ed.) *Geotechnical and Environmental Geophysics*. Tulsa: Society of Exploration Geophysicists.

Morrison HF and Becker A (1982) Analysis of airborne electromagnetic systems for mapping depth of seawater. University of California. *Final report ONR*, Contract N00014-82-M-0073.

Muffly G (1946) The airborne magnetometer. *Geophysics* 11: 321–334.

Nabighian MN (1979) Quasi-static transient response of a conducting half-space—An approximate representation. *Geophysics* 44: 1700–1705.

Oldenburg DW and Li Y (2005) Inversion for applied geophysics: A tutorial. In: Butler DK (ed.) *Near Surface Geophysics, Part 1: Concepts and Fundamentals*. Tulsa: Society of Exploration Geophysicists.

Olhoeft GR (1985) Low-frequency electrical properties. *Geophysics* 50: 2492–2503.

Oristaglio M and Spies B (eds.) (1999) *Three-Dimensional Electromagnetics*. Tulsa: Society of Exploration Geophysicists.

Palacky GJ (1988) Resistivity characteristics of geologic targets. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1. Tulsa: Society of Exploration Geophysicists.

Parasnis DS (1979) *Principles of Applied Geophysics*. London: Chapman and Hall.

Pfaffhuber A, Hendricks S, and Kvistedal Y (2012) Progressing from 1D to 2D and 3D near-surface airborne electromagnetic mapping with a multisensor, airborne sea-ice explorer. *Geophysics* 77: WB109–WB117.

Poddar M (1982) A rectangular loop source of current on a two-layered earth. *Geophysical Prospecting* 30: 101–114.

Poddar M (1983a) Electromagnetic sounding near a large square loop source of current. *Geophysical Prospecting* 31: 508–523.

Poddar M (1983b) A rectangular loop source of current on multilayered earth. *Geophysics* 48: 107–109.

Powell TL, Jansen J, and Taylor R (2003) Time domain electromagnetic induction survey of the deep sandstone aquifer in the Lake Winnebago Wisconsin area. In: *Symposium on the Application of Geophysics to Engineering and Environmental Problems 2003, San Antonio, TX. Environmental and Engineering Geophysical Society*, pp. 687–708.

Raiche AP (1987) Transient electromagnetic field computations for polygonal loops on a layered earth. *Geophysics* 52: 785–793.

Reid J, Worby A, Vrbancich J, and Munro A (2003) Shipborne electromagnetic measurements of Antarctic sea-ice thickness. *Geophysics* 68: 1537–1546.

Reitz JR, Milford FJ, and Christy RW (1979) *Foundations of Electromagnetic Theory*. Reading, MA: Addison-Wesley.

Ryu J, Morrison HF, and Ward SH (1970) Electromagnetic fields about a loop source of current. *Geophysics* 35: 862–896.

Ryu J, Morrison HF, and Ward SH (1972) Electromagnetic depth sounding experiment across Santa Clara Valley. *Geophysics* 37: 351–374.

Sengpiel KP (1988) Approximate inversion of airborne EM data from a multilayered ground. *Geophysical Prospecting* 36: 446–459.

Sonenshein RS (1997) Delineation and extent of saltwater intrusion in the Biscayne aquifer, eastern Dade County, Florida, 1995. U.S. Geological Survey. *Water-Resources Investigations Report 96-4285*.

Sørensen KI and Søndergaard VH (1999) Large-scale geophysical mapping and its application for ground water protection in urban areas. In: *Proceedings of the*

Symposium on the Application of Geophysics to Engineering and Environmental Problems, March 14–18, 1999, pp. 481–486. Oakland, CA: Environmental and Engineering Geophysical Society.
Spies B (1989) Depth of investigation in electromagnetic sounding methods. *Geophysics* 54: 872–888.
Spies BR and Eggers DE (1986) The use and misuse of apparent resistivity in electromagnetic methods. *Geophysics* 51: 1462–1471.
Spies BR and Frischknecht FC (1991) Electromagnetic sounding. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics—Applications*, vol. 2A. Tulsa: Society of Exploration Geophysicists.
Steuer A, Siemon B, and Eberle D (2008) Airborne and ground-based electromagnetic investigations of the freshwater potential in the tsunami-hit area Sigli, northern Sumatra. *Journal of Environmental and Engineering Geophysics* 13: 39–48.
Stratton JA (1941) *Electromagnetic Theory*. New York: McGraw-Hill Book Company.
Stumm W and Morgan JJ (1996) *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*. New York: John Wiley & Sons.
Sundberg K (1931) Principles of the Swedish geoelectrical methods. *Ergänzungshefte für Angewandte Geophysik* I.
Sundberg K and Hedstrom FH (1934) Structural investigations by electromagnetic methods. In: *Proceedings of the World Petroleum Congress, London*, pp. 102–110. Institute of Petroleum Technologists.
Tarantola A (1987) *Inverse Problem Theory—Methods for Data Fitting and Model Parameter Estimation*. Amsterdam: Elsevier.
Turekian KK (1976) *Oceans*. Englewood Cliffs, NJ: Prentice Hall.
Vanhala H, Lintinen P, and Ojala A (2009) Electrical resistivity study of permafrost on Ridnitšohkka fell in northwest Lapland, Finland. *Geophysica* 45: 103–118.
Vrbancich J (2012) Airborne electromagnetic bathymetry and estimation of bedrock topography in Broken Bay, Australia. *Geophysics* 77: WB3–WB17.
Vrbancich J, Fullagar P, and Smith R (2010) Testing the limits of AEM bathymetry with a floating TEM system. *Geophysics* 75: WA163–WA177.
Wait JR (1951) A conducting sphere in a time varying magnetic field. *Geophysics* 16: 666–672.
Wait JR (1953) A conducting permeable sphere in the presence of a coil carrying an oscillating current. *Canadian Journal of Physics* 31: 670–678.
Wait JR (1955) Mutual electromagnetic coupling of loops over a homogeneous ground. *Geophysics* 20: 630–637.
Wait JR (1982) *Geo-electromagnetism*. New York: Academic Press.
Ward SH (1967) Part C—The electromagnetic method. In: Hansen DA, Heinrichs WE Jr, Holmer RC, Macdougall RE, Rogers GR, Sumner JS, and Ward SH (eds.) *Mining Geophysics* Tulsa: Society of Exploration Geophysicists.
Ward SH and Hohmann GW (1988) Electromagnetic theory for geophysical applications. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1. Tulsa: Society of Exploration Geophysicists.
Waxman MH and Smits LJM (1968) Electrical conductivities of oil-bearing shaly sands. *Society of Petroleum Engineers Journal, Transactions of American Institute of Mining, Metallurgical and Petroleum Engineers* 243: 107–122.
Waxman MH and Thomas EJ (1974) Electrical conductivities of shaly sands. I. The relationship between hydrocarbon saturation and resistivity index. II. The temperature coefficient of electrical conductivity. *Journal of Petroleum Technology* 26: 213–225.
Wenner F (1912) The four-terminal conductor and the Thomson bridge. *Bulletin of the Bureau of Standards* 8: 559–610.
Wesley JP (1958a) Response of dyke to oscillating dipole. *Geophysics* 23: 128–133.
Wesley JP (1958b) Response of thin dyke to oscillating dipole. *Geophysics* 23: 134–143.
West GF and Macnae JC (1991) Physics of the electromagnetic induction exploration method. In: Misac NN (ed.) *Electromagnetic Methods in Applied Geophysics*, vol. 2. Tulsa: Society of Exploration Geophysicists, Part A.
Wilt M and Stark M (1982) A simple method for calculating apparent resistivity from electromagnetic sounding data. *Geophysics* 47: 1100–1105.
Winsauer WO, Shearin HM Jr., Mason PH, and Williams M (1952) Resistivity of brine-saturated sands in relation to pore geometry. *Bulletin of American Association of Petroleum Geologists* 36: 253–277.
Won IJ and Smits K (1987) Airborne electromagnetic measurements of electrical conductivity of seawater and bottom sediments over shallow ocean. *Marine Geotechnology* 7: 1–14.
Won IJ and Smits K (1990) Airborne electromagnetic bathymetry. In: Fitterman DV (ed.) *Developments and Applications of Modern Airborne Electromagnetic Surveys*: U.S. Geological Survey, U.S. Geological Survey Bulletin 1925.
Wright DL, Asch TH, Moulton CW, Irons TP, and Nabighian MN (2008) Effects of spatial data density, sensor noise and position errors on UXO and clutter target parameters from inversions of ALLTEM data. In: *Symposium on the Application of Geophysics to Engineering and Environmental Problems 2008*.
Wyllie MRJ and Gregory AR (1953) Formation factors of unconsolidated porous media: Influence of particle shape and effect of saturation. *Transactions of the American Institute of Mining, Metallurgical, and Petroleum Engineers* 198: 103–110.
Yang C-H, Tong L-T, and Huang C-F (1999) Combined application of dc and TEM to sea-water intrusion mapping. *Geophysics* 64: 417–425.
Zollinger R, Morrison H, Lazenby P, and Becker A (1987) Airborne electromagnetic bathymetry. *Geophysics* 52: 1127–1137.
Zonge K, Wynn J, and Urquhart S (2005) Resistivity, induced polarization, and complex resistivity. In: Butler DK (ed.) *Near Surface Geophysics, Part 1: Concepts and Fundamentals*. Tulsa: Society of Exploration Geophysicists.

This page intentionally left blank

11.11 Tools and Techniques: Magnetic Methods of Exploration – Principles and Algorithms

Y Li and M Nabighian, Colorado School of Mines, Golden, CO, USA

© 2015 Elsevier B.V. All rights reserved.

11.11.1	**Introduction**	336
11.11.2	**The Geomagnetic Field**	336
11.11.2.1	The Dipolar Field	337
11.11.2.2	The Nondipolar Field	337
11.11.2.3	Secular Variation	340
11.11.2.4	The International Geomagnetic Reference Field	340
11.11.2.5	The Time Variable Field	340
11.11.2.6	Local Magnetic Anomalies	340
11.11.2.7	Geomagnetic Reversals	340
11.11.3	**The Physics of Magnetism**	340
11.11.4	**Data Processing**	343
11.11.4.1	Upward Continuation	343
11.11.4.2	Downward Continuation	344
11.11.4.3	First Vertical Derivative	344
11.11.4.4	Second Vertical Derivative	345
11.11.4.5	Directional and Second-Order Derivatives	345
11.11.4.6	Reduction to the Pole	346
11.11.4.7	Pseudogravity Transformation	348
11.11.4.8	Analytic Signal	349
11.11.4.9	Amplitude Data	350
11.11.4.10	Source Parameter Imaging Method	351
11.11.4.11	Tilt Angle Method	352
11.11.4.12	Total Horizontal Gradient	353
11.11.4.13	Terracing	353
11.11.4.14	Matched Filtering	354
11.11.4.15	Werner Deconvolution	354
11.11.4.16	Euler Deconvolution	355
11.11.4.17	Wavelet Analyses	357
11.11.5	**Magnetic Forward Modeling**	357
11.11.5.1	Magnetic Anomaly	359
11.11.5.2	Forward Modeling of Distributed Sources	360
11.11.5.3	Full Solution of Induced Magnetization Based on Integral Equations	361
11.11.5.4	Solution of Induced Magnetization Based on Differential Equations	362
11.11.5.5	Numerical Example	363
11.11.6	**Equivalent Source Techniques**	363
11.11.6.1	Inverse Formulation of Equivalent Source Construction	364
11.11.6.2	Examples	367
11.11.6.2.1	A synthetic example	367
11.11.6.2.2	A field example	369
11.11.7	**Magnetic Susceptibility Inversion**	370
11.11.7.1	Susceptibility Inversion	371
11.11.7.2	Joint Inversion of Borehole and Surface Data	373
11.11.7.2.1	Sensitivity-based weighting function	374
11.11.7.2.2	Distance-based weighting function	374
11.11.7.3	Computational Aspects	375
11.11.7.3.1	Bound constraints through logarithmic barrier method	376
11.11.7.3.2	Wavelet compression-based acceleration	376
11.11.7.4	A Field Example	378
11.11.8	**Magnetization Inversion**	379
11.11.8.1	Direction Estimation	380
11.11.8.1.1	Helbig moment method	380
11.11.8.1.2	Wavelet multiscale edge method	381
11.11.8.1.3	Cross correlation methods	381
11.11.8.2	Amplitude Inversion	382

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00196-2

11.11.8.2.1 Synthetic example 384
11.11.8.3 Magnetization Vector Inversion 384
11.11.8.3.1 A synthetic example 386
11.11.8.4 A Field Example 386
11.11.9 Discussion 387
Acknowledgments 388
References 388

11.11.1 Introduction

The goal of this chapter is to provide information about the principles, understanding, algorithms, and applicability of the magnetic method to exploration and near-surface geophysical problems. Most targets of exploration, environmental, and archaeological type are located at shallow depths, and the magnetic method provides a cost-effective tool in their exploration.

The magnetic method is the oldest and one of the most widely used geophysical techniques in both mining and petroleum exploration. Although the magnetic properties of lodestones as a north-seeking rock were known in China around the first century AD, it was William Gilbert, physician to Queen Elizabeth I, who published in 1600 his landmark book *De Magnete* in which he summarized the work previously done by Chinese and European researchers and then, based on his experiments, concluded that the Earth was a magnet and that this was why the compass pointed north. In 1782, Pierre Simon, Marquis de Laplace, showed that a Newtonian potential obeys the now famous Laplace's equation, and in 1838, Carl Friedrich Gauss gave the first mathematical treatment to global-scale geomagnetic observations by using the mathematical formalism of spherical harmonic analysis.

The magnetic methods started being used in solving geologic problems in parallel with the advancements in magnetometers. The earliest magnetometer was the sundial compass that was used in 1640 in Sweden for iron ore prospecting. In the modern era, they were replaced by the balanced magnets that became the preferred tool for land-based magnetic surveys. They were quickly followed by fluxgate, proton precession, alkali vapor (cesium, rubidium, etc.), and Overhauser and SQUID magnetometers, which were soon applied also in both airborne and marine surveys. Recent developments include the magnetic vector and magnetic gradiometer instruments. Magnetic measurements can now also be made using satellites (e.g., LEO (low Earth orbit) and CHAMP (CHAllenging Minisatellite Payload)) to determine the broad magnetic variations in the lithosphere and thus enhance the geologic utility of near-surface surveys.

This chapter contains nine sections starting with this introduction and ending with a general discussion on covered topics.

Section 11.11.2 introduces the reader to the Earth's magnetic field, its dipolar and nondipolar components, its diurnal and secular variations, the International Geomagnetic Reference Field (IGRF), and magnetic reversals.

Section 11.11.3 introduces the basic ideas of magnetization, magnetic field, and magnetic induction, measurement units, and the concept of induced and remanent magnetizations.

Section 11.11.4 outlines the various methods used to process magnetic data for ease of interpretation, methods such as upward and downward continuation, first and second vertical derivatives, directional and second derivatives, reduction to the pole (RTP), pseudogravity (magnetic potential) transformation, analytic signal, source parameter imaging, tilt angle method, total horizontal gradient (THG), terracing, matched filtering, the Werner and Euler deconvolution, amplitude data, and wavelet analyses.

Section 11.11.5 presents various methods for the forward calculation of the anomaly due to either a single uniformly magnetized source or a distributed source with variable magnetization.

Section 11.11.6 presents a powerful alternative to the usual Fourier techniques for data processing by not requiring the data to be located on a horizontal plane and by not requiring the data to be gridded.

Section 11.11.7 outlines state-of-the-art methods of generalized physical property inversions for the magnetic susceptibility distribution in three-dimensional (3-D) space. The main assumption made is that all magnetization is induced with no remanent component. The basic inversion algorithm and various techniques for handling different computational issues are outlined.

Section 11.11.8 presents inversion techniques in the presence of both remanent magnetization and strong self-demagnetization effects. The methods presented include magnetization direction estimation and amplitude inversion.

Section 11.11.9 presents a final overview of covered topics and a look ahead.

Our wish is to present in a compact form the state of the art in using magnetic data for near-surface exploration. Modern methods of interpretation, based primarily on inversion, do complement the standard methods and help achieve a better quantitative interpretation than previously possible. We outline the basic mathematical concepts of each technique and illustrate the applicability of various methods with examples where appropriate. For the seasoned interpreters, we hope this chapter will help them when seeking answers to various magnetic interpretation problems. We also hope the materials will provide a concise explanation of the basic principle and algorithmic details for implementation and experimentation and, thereby, spur new research and development.

11.11.2 The Geomagnetic Field

The geomagnetic field of the Earth is composed of three main components:

(1) A large slowly varying component originating from inside the Earth

(2) A much smaller but rapidly varying component originating from outside the Earth due to currents flowing in the ionosphere
(3) Anomalies due to typical geophysical targets

To a first-order approximation, the Earth's magnetic field resembles the field of a magnetic dipole located at the center of the Earth and inclined about 11° with respect to the Earth's axis of rotation (Figure 1).

This field is oriented upward in the southern hemisphere and downward in the northern hemisphere and can best be described in a system of coordinates with the *x*-axis oriented towards geographic north, *y*-axis oriented towards geographic east, and *z*-axis oriented vertically down (Figure 2).

A magnetic field of strength $\vec{B}$ has an inclination I with respect to the horizontal XY plane. This inclination is positive in the northern hemisphere and negative in the southern hemisphere. The angle between the horizontal projection of $\vec{B}$ $\left(\vec{H}=\vec{B}\cos I\right)$ and the X-axis is the declination D that is positive east of the X-axis and negative otherwise. The $\vec{X}$, $\vec{Y}$, and $\vec{Z}$ components of the magnetic field can be calculated from the following:

$$\vec{X}=\vec{B}\cos I\cos D$$

$$\vec{Y}=\vec{B}\cos I\sin D$$

$$\vec{Z}=\vec{B}\sin I$$

Maps showing lines of equal declination D, inclination I, and field strength B are called isogonic, isoclinic, and isodynamic maps, respectively (Figure 3). The line of zero inclination is known as the magnetic equator. The Earth' field strength B varies between about 25 000 nT in the equatorial regions and about 70 000 nT in the polar regions, while the inclination I varies between −90° and +90°.

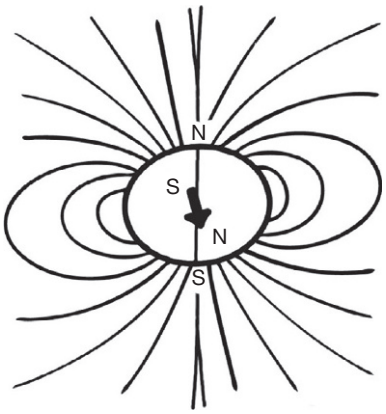


Figure 1 Earth's geomagnetic field.

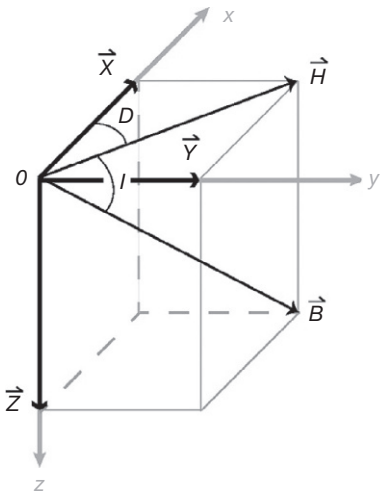


Figure 2 Earth's magnetic field parameters. *X*-axis is oriented along geographic north, *Y*-axis is oriented along geographic east, and the *Z*-axis is oriented vertically down.

11.11.2.1 The Dipolar Field

Spherical harmonic analysis, first applied by Gauss and refined later on, showed that 95% of the Earth's magnetic field originates from the interior of the Earth and is caused by convection currents of conducting material circulating in the conductive outer core and creating a geodynamo effect: the remaining 5% originates from outside sources related to the interaction between the *solar wind* and the Earth's internal magnetic field, with the ionosphere playing a significant role.

As seen earlier, to a first-order approximation, the Earth's magnetic field is *dipolar*, that is, it resembles the field of a magnetic dipole located at the center of the Earth and inclined about 11° with respect to the Earth's axis of rotation. The points on the Earth's surface where this dipole axis intersects the Earth are called the north and south *geomagnetic* poles. In 1990, the positive (north) pole was located at 79.1°S latitude and 108.9°E longitude, while the negative (south) pole was located at 79.1°N latitude and 71.1°W longitude. The points at which the Earth's magnetic field is normal to the surface of the Earth are different from the geomagnetic poles and are called the north and south *magnetic* poles. The north and south magnetic poles move continuously over time due to magnetic changes in the Earth's core. The north magnetic pole, where the magnetic field is pointing straight down (inclination of 90°) was located in 2005 near Ellesmere Island in northern Canada at 81.3°N and 110.8°W. The south magnetic pole, where the magnetic field is pointing straight up (inclination of −90°), was located in 2005 at 64.5°S and 137.8°E, just off the coast of Adélie Land in French Antarctica.

11.11.2.2 The Nondipolar Field

The departure between the Earth's magnetic field and the field of a dipole located at the center of the Earth is called the *nondipolar* field and represents < 10% of the total observed field. This field is composed of a few very large-scale magnetic anomalies of amplitudes of up to 20 000 nT (Figure 4). In principle, using spherical harmonic analysis, one can model the nondipolar field by using a small number of radial dipoles situated near the core–mantle interface although we should be careful in attaching any solid physical significance to such interpretation.

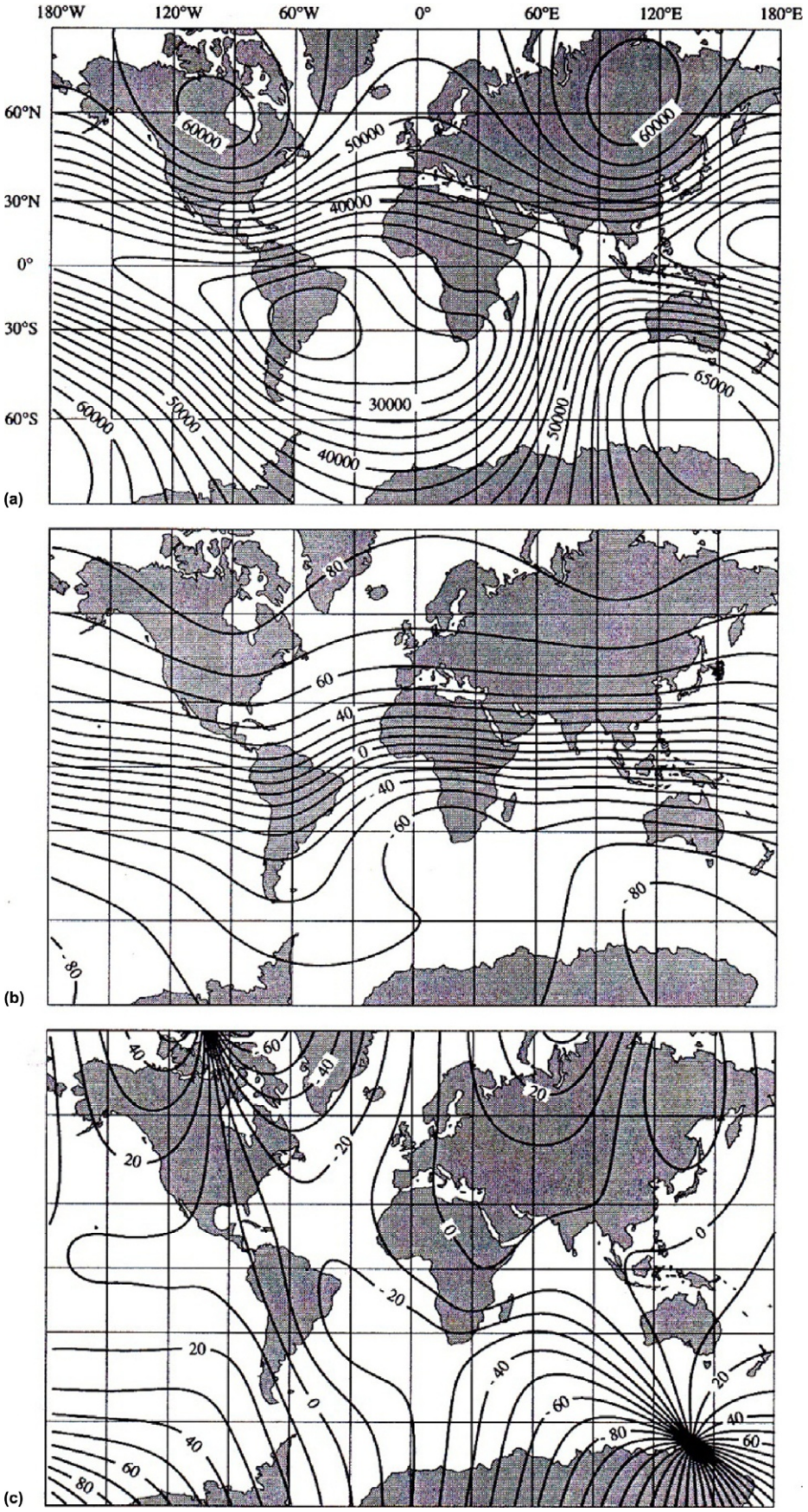


Figure 3 The magnetic field of the Earth for the year 1990. (a) Isodynamic map showing total intensity at a contour interval of 2500 nT. (b) Isoclinic map showing constant inclination at a contour interval of 10°. (c) Isogonic map showing constant declination at a contour interval of 10°. Reproduced from Blakely RJ (1996) Potential Theory in Gravity and Magnetic Applications. Cambridge: Cambridge University Press, 441 pp.

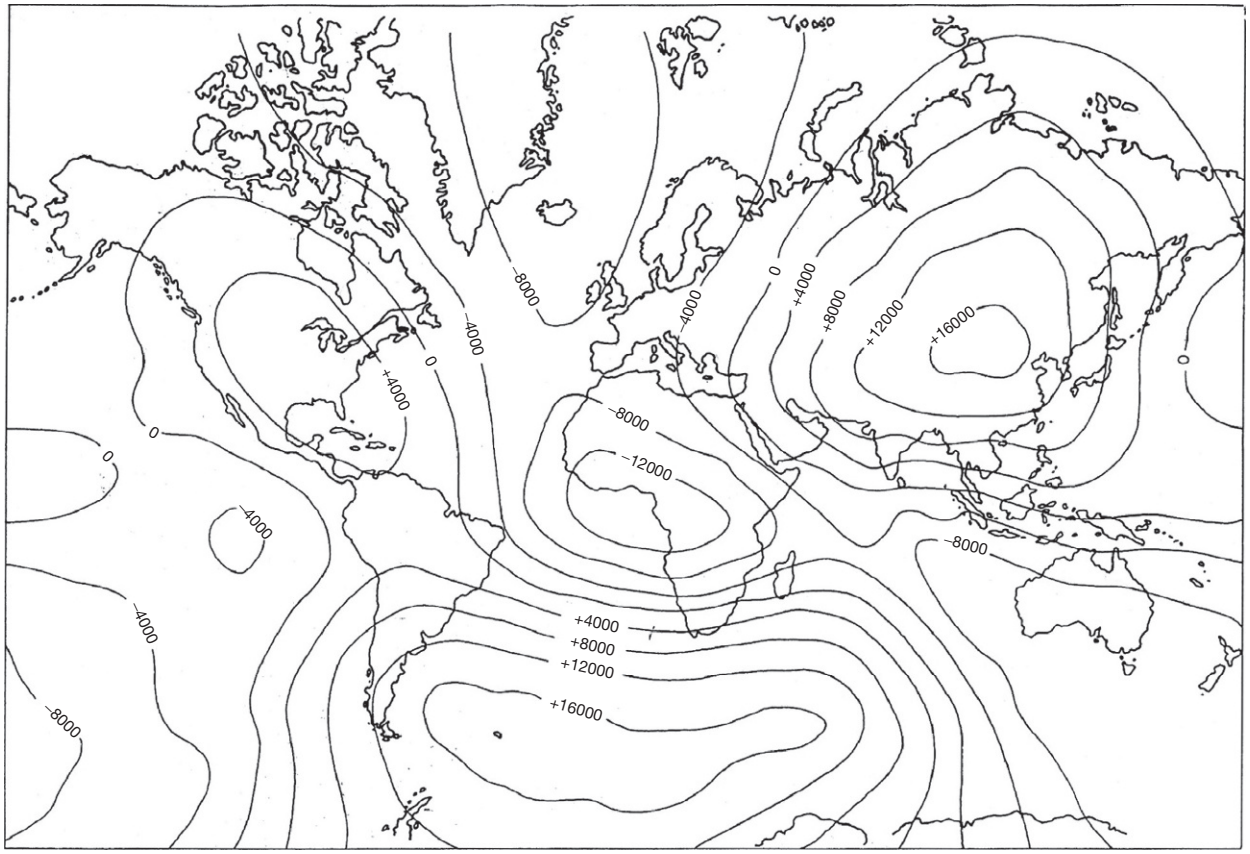


Figure 4 The nondipole part of the Earth's magnetic field for the year 1990 contoured at 4000 nT intervals. Reproduced from Blakely RJ (1996) Potential Theory in Gravity and Magnetic Applications. Cambridge: Cambridge University Press, 441 pp.

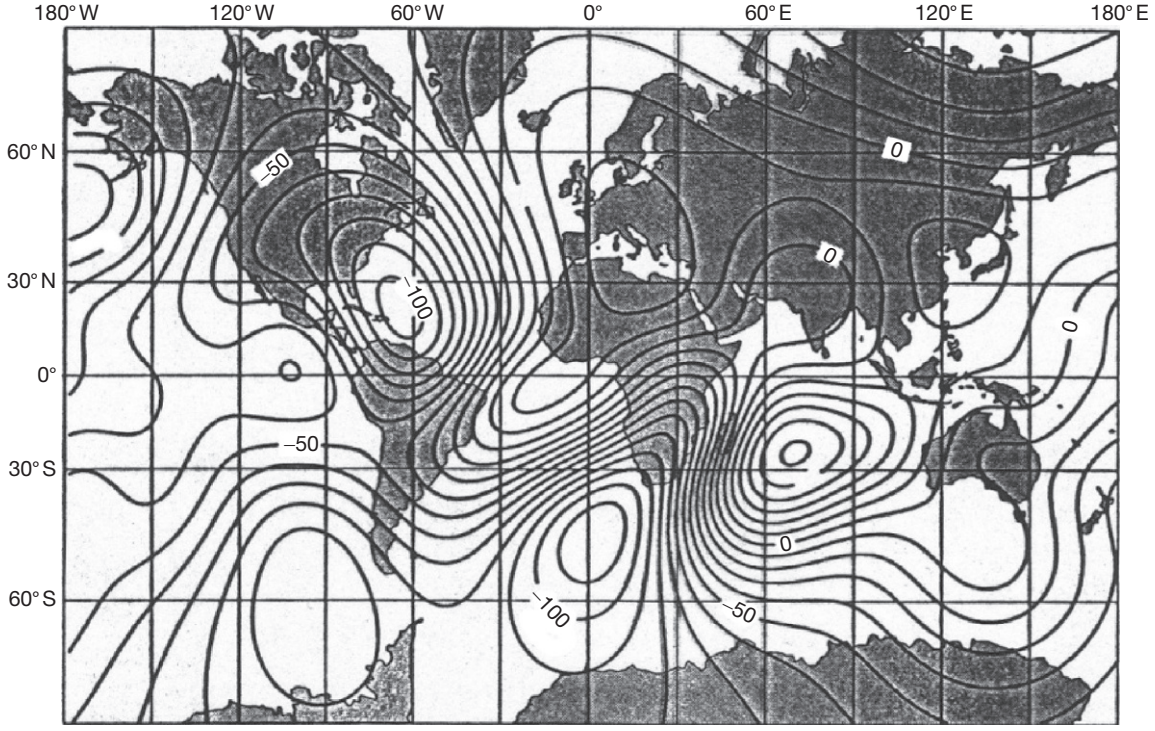


Figure 5 Secular variation for the period 1990–1991 in nT per year. Reproduced from Blakely RJ (1996) Potential Theory in Gravity and Magnetic Applications. Cambridge: Cambridge University Press, 441 pp.

11.11.2.3 Secular Variation

The dipolar and nondipolar components of the Earth's field are not constant but change slowly during a year (Figure 5). At a given location, this *secular* variation can reach amplitudes of over 100 nT $year^{-1}$ with a global average of about 50 nT $year^{-1}$. The secular variation pattern shows mostly a slow westward drift of about 0.2° longitude per year. Since the Earth's magnetic field originates in the Earth's liquid outer core, this westward drift implies that the core is rotating at a slightly slower rate than the crust or mantle: this observation was later on independently confirmed using seismic methods measuring earthquakes traveling through the Earth's core. There are some suggestions that the drift direction has not always been westward: some periods of eastward motion may have occurred in the last 3000 years.

11.11.2.4 The International Geomagnetic Reference Field

The Earth's geomagnetic field can be represented by using spherical harmonic analysis, and the resulting Gaussian coefficients are published every 5 years covering the next 5-year time period, which is known as an *epoch*. Currently, the IGRF is a spherical harmonic degree 13 model, and they allow the calculation of inclination, declination, and field strength for a given date and for a location specified by latitude/longitude and station elevation. In addition, these coefficients contain also a derivative term that allows the inclusion of secular variation for periods between the epochs. The published IGRF coefficients represent only the low-order terms that are assumed to model the field originating from within the core. Subtracting the IGRF field from the Earth's field will in principle result in anomalies caused by sources within the Earth's crust where the temperatures are less than the Curie temperatures of important magnetic minerals. This IGRF correction is an important first step in processing magnetic data.

11.11.2.5 The Time Variable Field

The Earth's magnetic field changes daily due to changes in the direction and strength of currents in the ionosphere. These *diurnal* variations are on average around 50 nT on quiet magnetic days. The maximum values at a given site occur around local noon time; the effects are larger as we approach lower latitudes, are stronger in the summer compared to winter, and are larger during daytime compared to nighttime. These diurnal variations are normally accounted for by using a base station. Periodically, one encounters days with *magnetic storms* when the Earth's field changes very rapidly over a short period of time reaching amplitudes of a few hundred nanoteslas followed by lower random amplitudes. During periods of such rapid high-amplitude changes, magnetic measurements cannot be carried out. Such days can however be known in advance by contacting a nearby magnetic observatory and inquiring about the probability of a magnetic storm in the next few days.

11.11.2.6 Local Magnetic Anomalies

Local magnetic anomalies of interest in geophysical exploration are usually located within the Earth's crust where temperatures are below the Curie temperatures of important magnetic minerals (see Moskowitz et al., this volume). The mantle is, in general, considered as nonmagnetic. These anomalies have lateral extents from kilometers for regional fields or petroleum targets to hundreds of meters for typical mining targets and meters for archaeological targets or unexploded ordinance. Their intensity can range from a few tenths of a nanotesla to a few thousand nanoteslas.

11.11.2.7 Geomagnetic Reversals

Periodically, the Earth's magnetic field reverses polarity by interchanging the north and south magnetic poles. The present direction of the magnetic field is called the '*normal*' field, while the field in the opposite direction is called the '*reversed*' field. The time span between reversals are called *chrons*, which have a time span of between 100 000 and 1 000 000 years with an average of 450 000 years (Figure 6). Most reversals are estimated to occur over a time span between 1000 and 10 000 years. The cause of these reversals is not fully understood. These reversals played an important role in proving the theory of seafloor spreading by observing the symmetrical pattern of magnetic anomalies around the Mid-Atlantic Ridge that shows alternate positive and negative polarities (Figure 7). The general term *magnetostratigraphy* is commonly used anytime magnetic reversals are used for timing a given geologic event.

11.11.3 The Physics of Magnetism

An electric current produces a magnetic field $\vec{H}$ that can be calculated using the Biot–Savart law. Such a field will align the magnetic dipoles $\vec{m}_i$ inside a magnetic body in the direction of the field, thus polarizing it. As a result, the body gets magnetized by *induction*. The intensity of magnetization $\vec{M}$ is defined as

$$\vec{M} = \frac{\sum \vec{m}_i}{V} \qquad [1]$$

where V is the volume of the rock.

The magnetization $\vec{M}$ is proportional to the magnetizing field $\vec{H}$:

$$\vec{M} = \kappa \vec{H} \qquad [2]$$

where the constant of proportionality κ is the magnetic susceptibility of the body and is a dimensionless quantity. By this definition, susceptibility is a measure of how susceptible a body is to be magnetized.

The combined effect of the magnetizing field $\vec{H}$ and the magnetization $\vec{M}$ yields the magnetic induction vector $\vec{B}$:

$$\vec{B} = \mu_0\left(\vec{H} + \vec{M}\right) = \mu_0(1+\kappa)\,\vec{H} = \mu\,\vec{H} \qquad [3]$$

where $\mu_0 = 4\pi\ 10^{-7}$ henry m^{-1} is the magnetic permeability in vacuum and $\mu = \mu_0(1+\kappa)$ is the magnetic permeability of the magnetic body itself, which is a measure of how much the force between two magnetic poles is modified when immersed in a magnetic media. The unit for $\vec{H}$ is ampere per meter and the unit for $\vec{B}$ is tesla = weber m^{-2}. In practice, a tesla is a very large unit, and in practice, one uses instead a nanotesla ($1\ nT = 10^{-9}$ T).

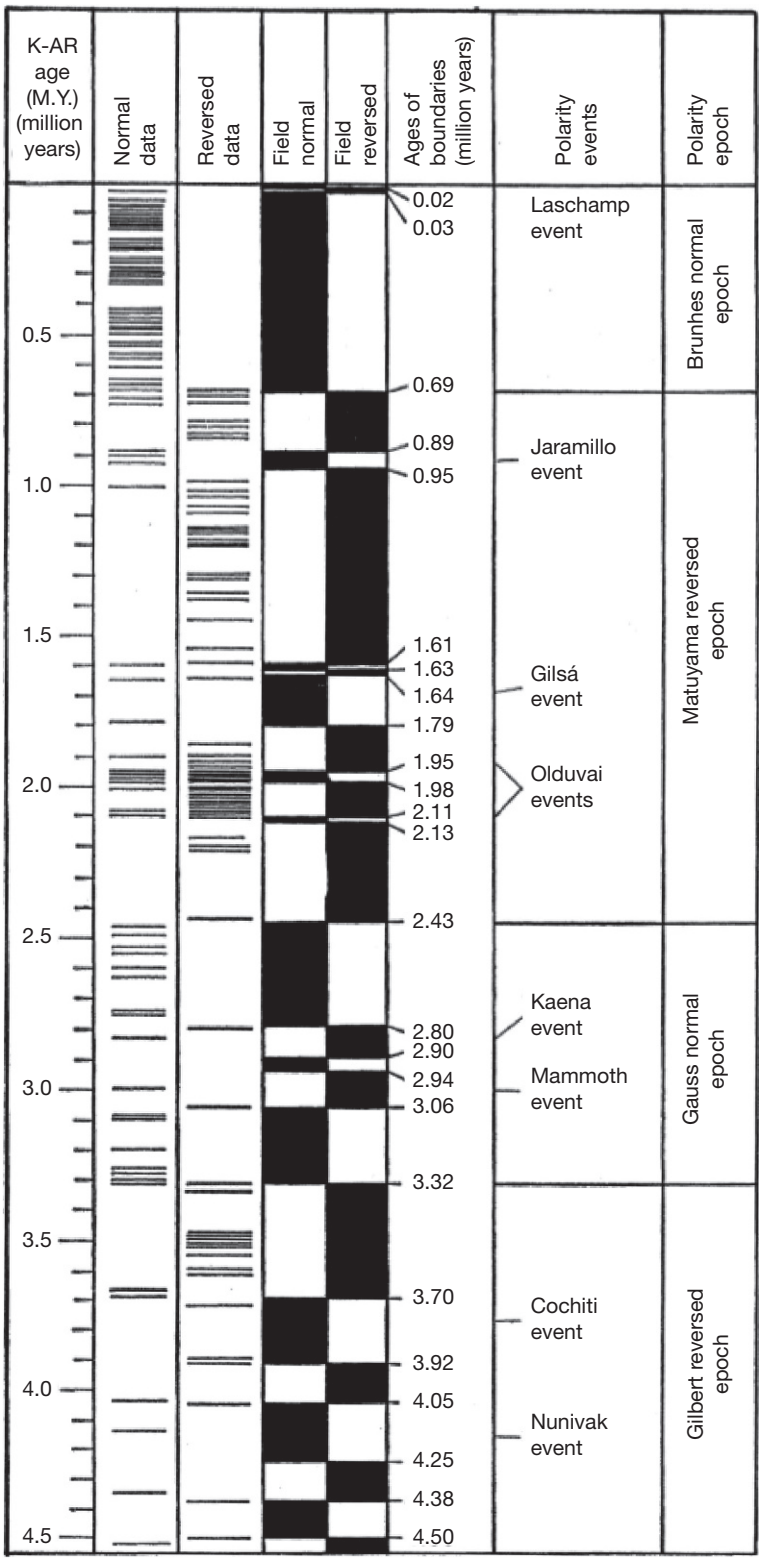


Figure 6 Main magnetic field reversals during the last 5 million years with the black lines indicating magnetization in the direction of the present-day field and white lines indicating reverse polarization. Reproduced from Cox A (1969) Geomagnetic reversals. *Science* 163: 237–245.

Oftentimes, in the absence of a field $\vec{H}$, many bodies still exhibit a measurable magnetization due to the presence of magnetic domains inside the body acquired at the time of either their formation (primary magnetization) or subsequent alteration (secondary magnetization). Such magnetization is known under the general term of *remanent* magnetization, which differs in both amplitude and direction from induced magnetization. There are a number of known processes under

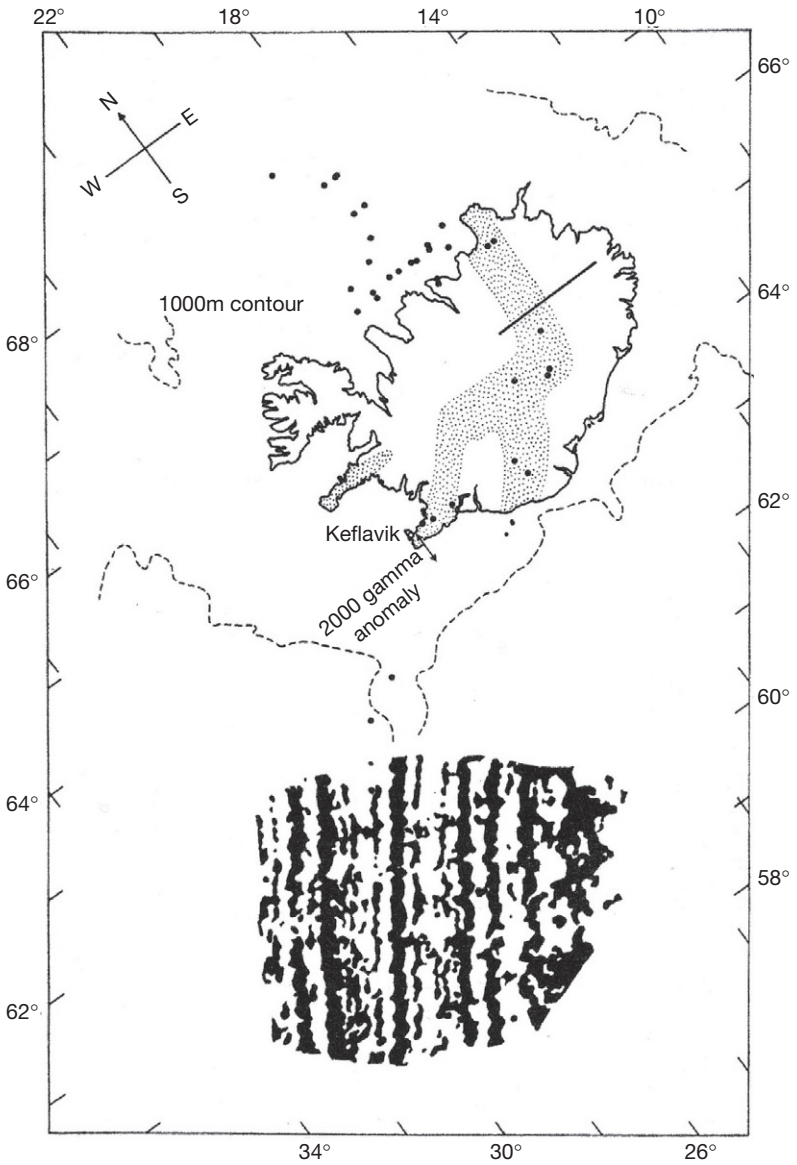


Figure 7 Pattern of symmetrical magnetic anomalies over the Mid-Atlantic Reykjanes Ridge south of Iceland. Reproduced from Heirtzler JR, LePichon X, and Baron JG (1966) Magnetic anomalies over the Reykjanes Ridge. *Deep Sea Research* 13: 427–432; Sheriff RE (1989) Geophysical Methods. New Jersey: Prentice Hall.

which such magnetization can take place with the notable ones being the thermal remanent magnetization (acquired by an igneous rock during cooling below Curie temperature) and depositional remanent magnetization (acquired when sediments while depositing towards the ocean bottom tend to align in the direction of the existing magnetic field). Remanent magnetization can be also of natural, isothermal, chemical, thermochemical, viscous, and anhysteretic origin (Table 1). The ratio between the intensity of remanent magnetization and that of induced magnetization is known as the Königsberger ratio Q, which can reach values as high as 100 for some rocks.

The magnetic minerals in various rocks can in general be separated into five main categories:

(a) *Diamagnetic* minerals have a small negative magnetic susceptibility due to the fact that the negatively charged electrons subjected to a magnetizing field $\vec{H}$ alter their rotation rate to oppose the field. All magnetic materials exhibit this property, which is weakly affected by temperature, but for most of the rocks, this effect is overwhelmed by the normal magnetization effects. Notable diamagnetic rocks are rock salt, anhydrite, quartz, marble, and graphite.

(b) *Paramagnetic* rocks have positive magnetic susceptibilities and their effect decreases with absolute temperature. In paramagnetic rocks, the magnetic moments align partially with the magnetizing field $\vec{H}$, thus producing a net magnetization in the direction of the applied field. Thermal effects destroy this alignment and paramagnetic effects disappear in the absence of a magnetizing field.

(c) *Ferromagnetic* rocks (Figure 8(a)) are represented by minerals such as iron, nickel, and cobalt and are characterized by large magnetic domains aligned all in the same

Table 1 Types of natural remanent magnetization (after Olhoeft, G., personal communication)

ARM (anhysteretic remanent magnetization)	Produced by an external field in the presence of a decaying-amplitude alternating field
CRM (chemical (or crystallization) remanent magnetization)	Produced by a chemical change (such as oxidation–reduction) after deposition or crystallization
DRM (depositional (or detrital) remanent magnetization)	Produced by small magnetic grains being oriented by the Earth's magnetic field while being deposited
IRM (isothermal remanent magnetization)	Produced by the removal of an external field at constant temperature
NRM (natural remanent magnetization)	Natural or *in situ* remanence of a sample (however produced)
PRM (pressure (or piezoremanent) remanent magnetization)	Produced by mechanic defamation while in a magnetic field
TRM (thermal remanent magnetization)	Produced by cooling material from above to below its Curie temperature in a magnetic field
VRM (viscous remanent magnetization)	Produced over long periods of time exposure to a constant external magnetic field

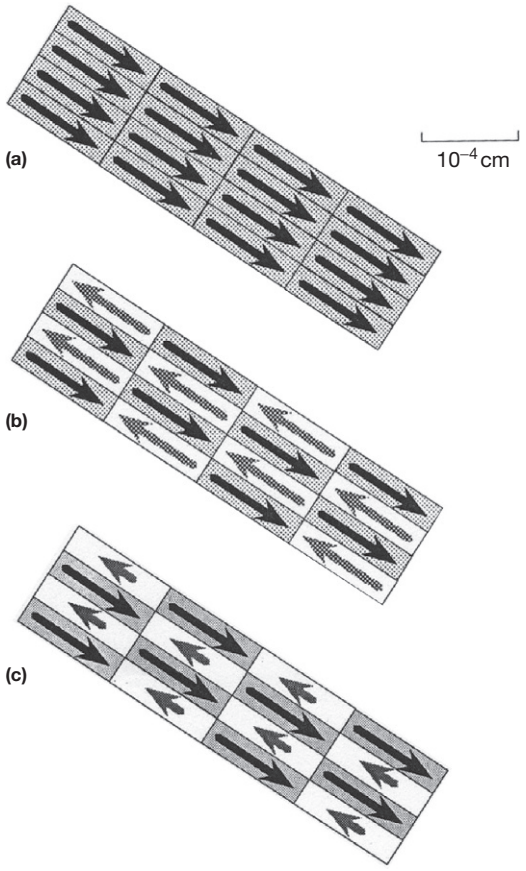


Figure 8 Schematic representation of magnetic domains in (a) ferromagnetic, (b) antiferromagnetic, and (c) ferrimagnetic minerals. Reproduced from Burger HR (1992) Exploration Geophysics of the Shallow Subsurface. New Jersey: Prentice Hall.

direction, resulting in very large magnetic susceptibilities. Ferromagnetic materials do not exist inside the Earth but are present in meteorites. Ferromagnetism also decreases with temperature and disappears at the Curie temperature.

(d) *Antiferromagnetic* rocks (Figure 8(b)) have practically an equal number of domains in both directions resulting in very small susceptibilities. Hematite is the most common antiferromagnetic mineral.

(e) *Ferrimagnetic* rocks (Figure 8(c)) exhibit some domains that are oriented in the opposite direction to the general domain orientation but still do exhibit a net magnetization. Magnetite, ilmenite, titanomagnetite, pyrrhotite, and various iron oxides are examples of ferrimagnetic rocks.

11.11.4 Data Processing

Since there is a strong relationship between the dominant wavelengths of a magnetic anomaly and the size, shape, and depth of the causative body, it becomes apparent that spectral analysis techniques will play an important role in providing further insight in interpreting magnetic data and understanding the nature of magnetic sources. As a result, Fourier transform techniques have become an effective and efficient tool and have been used ubiquitously since the advent of digital computing. In addition, the fact that magnetic fields are related to the source through a convolution with Green's function also makes the Fourier transform a powerful tool to process magnetic data based on the convolution theorem. Wavelet transform techniques have also seen increased use because of their ability to extract location-dependent spectral properties. The reader is cautioned though that all these techniques depend on certain assumptions and prerequisites not always easily achievable in practice, such as the requirement that data be collected on a planar surface in order to properly apply the Fourier transform. As a result, such processed maps should always be compared with the original magnetic data to make sure that no fictitious anomalies were created by such processing.

11.11.4.1 Upward Continuation

Due to the $1/r^3$ dependence of the magnetic field, a shallow target contains more energy in the shorter wavelength components than a deeper target, and its anomaly amplitude decreases much faster with target depth. Thus, measuring the magnetic field at a certain height above the original observational surface would lead to a stronger attenuation of shallower anomalies with only minor attenuation of deeper ones. Fortunately, we do not have to actually carry out measurements at a higher elevation for this purpose since we can reliably calculate the field analytically at a higher elevation. (When the initial magnetic survey covers only a small area, one cannot reliably calculate the upward-continued field at relatively high elevations, and

a higher altitude survey becomes necessary.) This upward continuation operation has many useful applications in practice. For instance, a deep-seated intrusive target can be easier interpreted by greatly reducing the high-frequency 'noise' created by shallower sources. Also, if we have two contingent airborne magnetic surveys carried out at different elevations (e.g., a marine survey at sea level and an airborne survey), we can use upward continuation to bring the lower-elevation survey to the level of the higher-elevation one, thus facilitating the merging of the two data sets.

The earliest attempts at upward continuation were done in the space domain by deriving a set of weights, which, when convolved with field data, yielded approximately the desired transform (Byerly, 1965; Fuller, 1967; Henderson, 1960; Peters, 1949). The space-domain operators were soon replaced by frequency-domain operators by recognizing the utility of using Fourier transform techniques in performing analytic continuations (Dean, 1958).

The classic derivation of the upward continuation operator was given by Henderson (1970) and was based on Green's third identity. Henderson showed that for a system of coordinates with the z-axis oriented upward, the total-field anomaly $M(x, y, z)$ at a height z can be obtained from the anomaly on the surface of the Earth $M(x,y,0)$ by integrating it over the $z=0$ plane:

$$M(x, y, z) = \frac{z}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{M(\alpha, \beta)\, d\alpha\, d\beta}{\left[(x-\alpha)^2 + (y-\beta)^2 + z^2\right]^{\frac{3}{2}}} \qquad [4]$$

Similar expressions are valid also for any measured magnetic component and also for the magnetic potential. In what follows, we will use the general notation $M(x,y,z)$ to signify any magnetic quantity unless specified otherwise.

Since the expression mentioned earlier is a convolution integral, its Fourier transform can be written in the form

$$\widetilde{M}(x, y, z) = \widetilde{M}(x, y, 0)\mathrm{F}\left[\frac{1}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}\right]$$
$$= \widetilde{M}(x, y, 0)e^{-z\sqrt{u^2+v^2}} \quad \text{for} \quad z \geq 0 \qquad [5]$$

where F and $\widetilde{M}(x, y, z)$ signify Fourier transform and u and v are the wave numbers in the x- and y-direction, respectively. Letting $\rho = \sqrt{u^2 + v^2}$ be the radial wave number, the expression noted earlier can be written in compact form as

$$\widetilde{M}(x, y, z) = \widetilde{M}(x, y, 0)e^{-z\rho} \quad \text{for} \quad z \geq 0 \qquad [6]$$

Thus, to continue the magnetic data upward a distance z, one first takes the Fourier transform of the data, and, after filtering (weighting) with the function $e^{-z\rho}$, one uses the inverse Fourier transform to obtain the upward-continued values. This procedure is common to all filtering operations described in this section: The data are first Fourier-transformed, and, after applying the proper filter, the data are inverse Fourier-transformed to obtain the desired result in the space domain.

The upward continuation filter is remarkably stable numerically; it is bell-shaped and has values ranging only from 0 (at infinite frequency) to 1 (at zero frequency). From eqn [6], it can be easily seen that upward continuation is a low-pass operation designed to reduce the high-frequency component of the data spectrum.

11.11.4.2 Downward Continuation

In a manner similar to the previous development, one can also downward continue the magnetic data to a lower elevation, as long as there are no sources present along the way. Such an operation will be useful in better identifying shallower sources since we are simulating taking measurements closer to these sources. Assuming that the measurements were made at an elevation z_1 and we would like to downward continue at a lower elevation zs_2, we immediately obtain

$$\widetilde{M}(x, y, z_2) = \widetilde{M}(x, y, z_1)e^{h\rho} \qquad [7]$$

where $h = z_1 - z_2$ and $z_2 < z_1$.

In contrast with the upward continuation filter, the downward continuation filter $e^{h\rho}$ is not stable, having values ranging from 1 (at zero frequency) to infinity (at infinite frequency). To avoid the instability of this filter at higher frequencies, it is customary to taper the filter gradually to zero at these higher frequencies with the taper cutoff frequency chosen based on the noise level and downward continuation distance. Although this approach will reduce the instability at high frequencies, it will result in less accurate data. Regardless, downward continuation is often used in many practical applications when shallower sources are the target.

A better behaved filter for the downward continuation operator can be obtained by making use of the Tikhonov regularization:

$$\frac{e^{h\rho}}{1 + \alpha\rho^2 e^{2h\rho}} \qquad [8]$$

where $\rho = \sqrt{u^2 + v^2}$ and α is a regularizing parameter to be chosen according to the amount of data noise (note that for $\alpha = 0$, eqn [8] reduces to eqn [7]). The factor ρ^2 implies that the downward-continued field is the flattest in the space domain (minimizing the square of the first-order derivative in both horizontal directions). Alternatively, one can incorporate the spectral decay of the magnetic data and determine a data-adaptive filter of the form (Li et al., 2013)

$$\frac{e^{h\rho}}{1 + \alpha P_0(\rho) e^{2h\rho}} \qquad [9]$$

where $P_0(\rho)$ is the radial power spectrum of the downward-continued data, which can be obtained by first estimating a parametric form of the radial power spectrum of the data using ensemble analysis (Spector and Grant, 1970) and then changing the ensemble depths to those below the downward continuation plane. The proof of these two expressions (eqns [8] and [9]) is similar to that for the stable RTP in the succeeding text (see Tikhonov and Arsenin, 1979, for the theory and Li et al., 2013, for specific details). An example of downward continuation is shown in **Figure 9** as applied to unexploded ordnance (UXO) detection.

11.11.4.3 First Vertical Derivative

Since shallower sources attenuate faster with height in comparison with deeper sources, we can use the first derivative with respect to z to again help separate deeper sources from shallower ones. After taking the z derivative of eqn [6] and then setting $z = 0$, we obtain

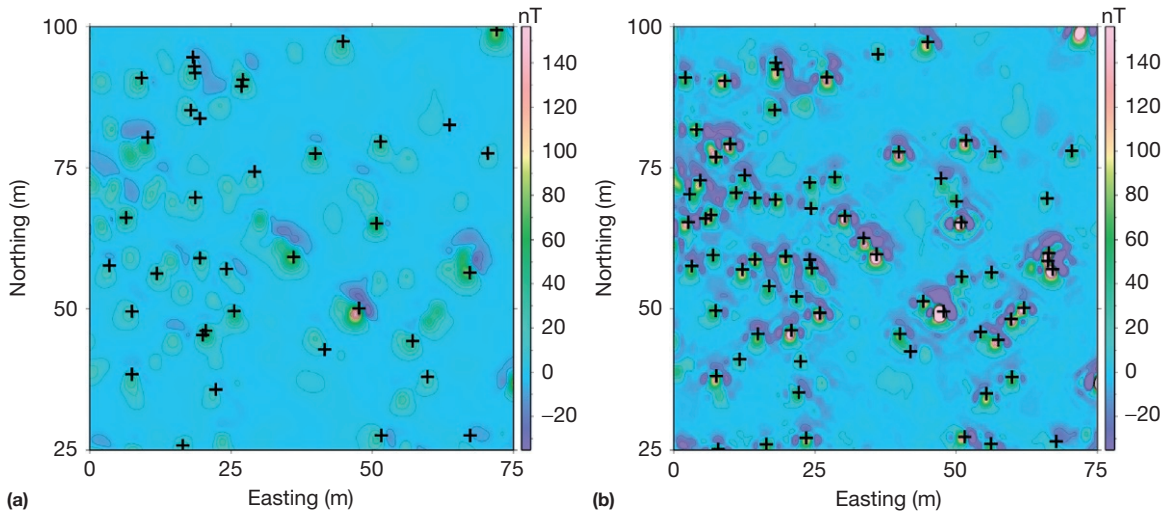


Figure 9 Example of downward continuation using the stabilized filter in eqn [9]. (a) The original airborne magnetic data acquired at a height of 1.25 m using a helicopter platform (color contours) and the corresponding unexploded ordnance (UXO) detection results. The grid spacing is 0.25 m. (b) Stable downward-continued data at 1 m (four grid spacing) below the original observation plane and the improved UXO detection results. Reproduced from Li Y, Devriese SGR, Krahenbuhl RA, and Davis K (2013) Enhancement of magnetic data by stable downward continuation for UXO application. *IEEE Transactions on Geoscience and Remote Sensing* 51: 3605–3614.

$$\frac{\partial}{\partial z}\widetilde{M}(x,y,0)=\rho\widetilde{M}(x,y,0)\quad\text{for}\quad z=0 \qquad [10]$$

The first vertical derivative filter $\rho=\sqrt{u^2+v^2}$ is not well behaved having values ranging from 0 (at zero frequency) to infinity (at infinite frequency). As before, a tapering of this filter is applied before doing the actual calculations. Using this filter, the higher frequencies due to shallower targets will be amplified with the result that adjacent anomalies can be more easily separated.

Both a more stable and a more accurate first vertical derivative filter can be calculated using Hilbert transforms (Nabighian, 1984). Starting from the identity

$$\sqrt{u^2+v^2}=iu\frac{-iu}{\sqrt{u^2+v^2}}+iv\frac{-iv}{\sqrt{u^2+v^2}} \qquad [11]$$

we can write (by recognizing the meaning of various terms):

$$\frac{\mathrm{d}\widetilde{M}}{\mathrm{d}z}=\frac{-iu}{\sqrt{u^2+v^2}}\frac{\mathrm{d}\widetilde{M}}{\mathrm{d}x}+\frac{-iv}{\sqrt{u^2+v^2}}\frac{\mathrm{d}\widetilde{M}}{\mathrm{d}y} \qquad [12]$$

The x and y derivatives of the magnetic field can be reliably computed in space domain, and their Fourier transforms are then multiplied by filters that only vary with frequency from -1 to 1, thus avoiding any need for filter tapering and thus increasing the accuracy of first vertical derivative calculations. The two finite filters mentioned earlier are the x- and y-components of the Hilbert transform operator (Nabighian, 1984).

11.11.4.4 Second Vertical Derivative

A second derivative is a measure of curvature. Since the magnetic anomaly of a shallow target decreases with height faster than the anomaly of a deeper target, it follows that the curvature of a shallow target will be larger and this can again be used to emphasize shallower targets at the expense of deeper ones. Taking once more the z derivative of eqn [6], we obtain for $z=0$

$$\frac{\mathrm{d}^2\widetilde{M}}{\mathrm{d}z^2}=\left(u^2+v^2\right)\widetilde{M}=\rho^2\widetilde{M} \qquad [13]$$

The second vertical derivative filter $\rho^2=u^2+v^2$ is again not well behaved having values ranging from 0 (at zero frequency) to infinity (at infinite frequency). As for the case of the first derivative filter, a tapering of this filter is often required before doing the actual calculations. Fortunately, the second derivative can be reliably calculated in the space domain using Laplace's equation to obtain

$$\frac{\partial^2 M}{\partial z^2}=-\left(\frac{\partial^2 M}{\partial x^2}+\frac{\partial^2 M}{\partial y^2}\right) \qquad [14]$$

In the expression mentioned earlier, the second derivatives with respect to x and y can easily be calculated in the space domain using finite difference expressions.

In the early days, the second derivative filter was used extensively when interpreting magnetic data since, among others, its zero contour helped define approximately the edges of these sources (Vacquier et al., 1951). In addition, the second derivative filter also helps in detecting and examining noise in data sets.

11.11.4.5 Directional and Second-Order Derivatives

Starting from the vertical component Z that can be obtained via a half RTP operation (see eqn [33]), we can calculate any other magnetic component, including the tensor components measured in magnetic gradiometry, by simple manipulations in the frequency domain. The derivations are greatly simplified by first calculating the magnetic potential V and then taking the appropriate derivatives. Since the vertical component is defined as

$$Z = \frac{\partial V}{\partial z} = V_z \qquad [15]$$

by using the first derivative filter defined earlier, it follows that

$$\widetilde{Z} = \sqrt{u^2 + v^2}\ \widetilde{V} = \rho \widetilde{V} \qquad [16]$$

or

$$\widetilde{V} = \frac{\widetilde{Z}}{\sqrt{u^2 + v^2}} = \frac{\widetilde{Z}}{\rho} \qquad [17]$$

which yields the scalar potential V.

From eqn [17], we obtain the x- and y-components of the magnetic field:

$$\begin{aligned} \widetilde{X} &= iu \frac{\widetilde{Z}}{\sqrt{u^2 + v^2}} = iu \frac{\widetilde{Z}}{\rho} \\ \widetilde{Y} &= iv \frac{\widetilde{Z}}{\sqrt{u^2 + v^2}} = iv \frac{\widetilde{Z}}{\rho} \end{aligned} \qquad [18]$$

The tensor components can be immediately obtained by differentiating the X-, Y-, and Z-components in the x-, y-, and z-direction, respectively. We thus obtain

$$\begin{aligned} \widetilde{V}_{xx} &= -u^2 \frac{\widetilde{Z}}{\sqrt{u^2 + v^2}} = -u^2 \frac{\widetilde{Z}}{\rho} \\ \widetilde{V}_{xy} &= -uv \frac{\widetilde{Z}}{\sqrt{u^2 + v^2}} = -uv \frac{\widetilde{Z}}{\rho} \\ \widetilde{V}_{xz} &= iu\widetilde{Z} \end{aligned} \qquad [19]$$

and similar expressions for the remaining components.

The expressions mentioned earlier can be written in matrix form as

$$\begin{pmatrix} \widetilde{V}_x \\ \widetilde{V}_y \\ \widetilde{V}_z \end{pmatrix} = \begin{pmatrix} \frac{iu}{\rho} \\ \frac{iv}{\rho} \\ 1 \end{pmatrix} \widetilde{Z} \qquad [20]$$

and

$$\begin{pmatrix} \widetilde{V}_{xx} & \widetilde{V}_{xy} & \widetilde{V}_{xz} \\ \widetilde{V}_{yx} & \widetilde{V}_{yy} & \widetilde{V}_{yz} \\ \widetilde{V}_{zx} & \widetilde{V}_{zy} & \widetilde{V}_{zz} \end{pmatrix} = \begin{pmatrix} \frac{-u^2}{\rho} & \frac{-uv}{\rho} & iu \\ \frac{-uv}{\rho} & \frac{-v^2}{\rho} & iv \\ iu & iv & \rho \end{pmatrix} \widetilde{Z} \qquad [21]$$

The filters required to carry out the calculations in eqns [20] and [21] are again not well behaved and special care must be taken when applying them.

11.11.4.6 Reduction to the Pole

The magnetic anomaly pattern over a given target is quite complicated and strongly dependent on magnetization direction and the measured magnetic component. Baranov (1957) and Baranov and Naudy (1964) proposed a simple method of transforming total-field anomalies into simpler anomalies, thus eliminating the distortion due to the magnetic field inclination. This RTP technique transforms existing magnetic data to data that would have been measured at the magnetic pole where the anomalies will be more symmetrical and better located over the target.

For the general case, the total-field magnetic anomaly T can be calculated at an external point $\vec{r}$ from

$$T\left(\vec{r}\right) = M \frac{\partial^2}{\partial \alpha \partial k} \int_V \frac{\mathrm{d}v}{\left|\vec{r} - \vec{r}_0\right|} = \frac{\partial^2 V}{\partial \alpha \partial k} \qquad [22]$$

where α is the direction of total magnetization (induced plus remanent components), k is the direction of measurement, V is the potential function, and $\vec{r}_0$ is the integration point inside the target. In the space domain, we can thus write for the total-field anomaly:

$$T\left(\vec{r}\right) = \left(l\frac{\partial}{\partial x} + m\frac{\partial}{\partial y} + n\frac{\partial}{\partial z}\right)\left(L\frac{\partial}{\partial x} + M\frac{\partial}{\partial y} + N\frac{\partial}{\partial z}\right)V \qquad [23]$$

where l, m, and n and L, M, and N are the direction cosines of the measured component and total magnetization directions, respectively.

In the frequency domain, the expression mentioned earlier becomes

$$\begin{aligned} T(u, v) &= (liu + miv + n\rho)(Liu + Miv + N\rho)V(u, v) \\ &= \rho^2 \left(l\frac{iu}{\rho} + m\frac{iv}{\rho} + n\right)\left(L\frac{iu}{\rho} + M\frac{iv}{\rho} + N\right)V(u, v) \end{aligned} \qquad [24]$$

which leads to the RTP filter:

$$\frac{1}{\left(l\frac{iu}{\rho} + m\frac{iv}{\rho} + n\right) \cdot \left(L\frac{iu}{\rho} + M\frac{iv}{\rho} + N\right)} \qquad [25]$$

An example of data reduced to the pole and their ability to directly map geology is illustrated in **Figure 10** over an extended area. **Figure 10(a)** shows the regional geology of the area, while **Figure 10(b)** shows the total-field anomaly map over the same area. Because of the low magnetic latitude, the total-field magnetic data have little correspondence with the general structural trends shown in the geology map in **Figure 10(a)**: there is however a good agreement between known geology and magnetic data once the data were reduced to the pole (**Figure 10(c)**).

At low magnetic latitudes, however, the RTP operator becomes unbounded along the direction of magnetic declination, and, by amplifying the noise in this direction, the resultant RTP field is dominated by linear features aligned along the declination direction. Baranov and Naudy (1964) developed the first procedure to calculate the RTP field based on integrating directly eqn [22] in the space domain but did not address the instability at low magnetic inclinations. Bhattacharyya (1965) expanded the magnetic field in a double Fourier series and carried out the double integration in eqn [22] analytically but again did not overcome the instability problem at low latitudes. Gunn (1972, 1995) designed Wiener filters in the space domain by determining filter coefficients that transform a known input model at the survey location to a desired output at the pole but again did not overcome the instability problem at low latitudes.

Faced with this difficulty, much effort has been spent on formulating stable approximations of the RTP operator.

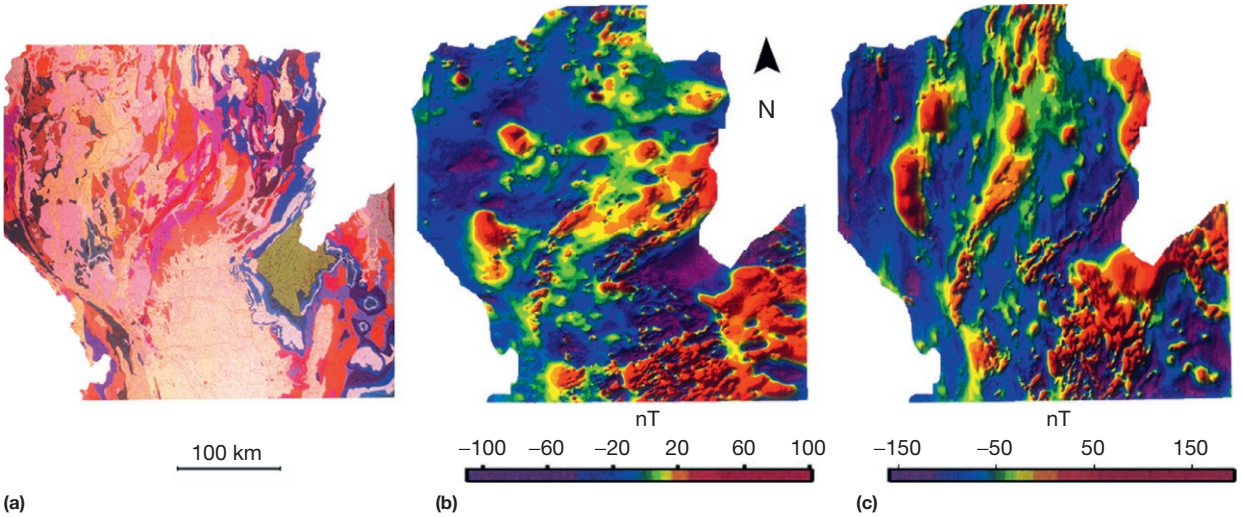


Figure 10 A large-scale example illustrating the ability of RTP data in directly mapping geology. (a) Simplified geologic map, (b) total-field magnetic anomaly, (c) RTP data.

Grant et al. (1972) introduced the concept of 'pseudo-inclination' in tackling this problem but never published their results. Pearson and Skinner (1982) proposed a whitening approach that strongly reduced the peak amplitude of the RTP filter, thus reducing noise. Silva (1986) proposed using equivalent sources that led to a stable but computationally intensive technique at that time, while Mendonca and Silva (1993) proposed using a stable truncated series approximation.

Most approaches focused upon the stabilization of the operator without regard to the actual noise in the data, and thus, postreduction processing was sometimes required to suppress the noise. Hansen and Pawlowski (1989), however, developed an elegant method in which a Wiener filtering technique was used to form a regularized RTP operator. The important contribution of their work was the stabilization of the RTP operator that was carried out in accordance with the noise in the observed field in the form of an estimated noise power spectrum. Since this approach, and others in the literature, is based upon filtering, little emphasis was placed upon being able to reproduce the observed data. Keating and Zerbo (1996) emphasized the need for better reproduction of the data as a factor for improved RTP results.

Li and Oldenburg (1998b, 2000a) proposed a technique that attempts to find the RTP field under the general framework of an inverse formulation, with the RTP field constructed by solving an inverse problem in which a global objective function is minimized subject to fitting the observed data. Assume that we have a set of observed total-field magnetic data acquired on a planar observation surface and produced by a magnetization in the direction $\hat{B}_0$. As is well known, the observed total-field anomaly is well approximated by the projection of the actual anomaly in the direction of the Earth's field. Let $\widetilde{T}(u,v)$ be the Fourier transform of the observed magnetic data and $\widetilde{R}(u,v)$ be the Fourier transform of the RTP field that would be observed at the same location. They defined a projection operator that relates the total field to the RTP field:

$$\widetilde{T}(u,v)=\frac{\left(\vec{K}\cdot\hat{B}_0\right)^2}{u^2+v^2}\widetilde{R}(u,v) \qquad [26]$$

where $\vec{K}=\left(iu, iv, \sqrt{u^2+v^2}\right)$ is the wave number vector. This projection can be considered as a forward mapping operator that calculates the magnetic anomaly at an arbitrary latitude from the true field at the pole. This operator is stable and unique. The RTP operator in eqn [25] is the direct inverse for the projection operator and, like many mathematical inverse problems in geophysics, is unstable. Based upon this understanding, it is logical to use the well-established inverse methodology in geophysics to deal with this problem. Using such an approach, the Fourier transform of the total-field anomaly would be the data, whereas that of the RTP field would be the model to be recovered.

Given the use of fast Fourier transform (FFT) in practice, the quantities in eqn [26] are defined at discrete points (u_i, v_i) over the rectangular grid. Adopting a commonly used ordering, we can collect the quantities into matrix–vector notations:

$$\mathbf{G}\,\vec{m}=\vec{d} \qquad [27]$$

where

$$\begin{aligned}\vec{d}&=\left(\ldots,\widetilde{T}(u_i,v_i),\ldots\right)^T\\ \vec{m}&=\left(\ldots,\widetilde{R}(u_i,v_i),\ldots\right)^T\\ \mathbf{G}&=diag\{G(u_i,v_i)\}\end{aligned} \qquad [28]$$

and the superscript T denotes transposition. In this form, it is easier to see that the matrix G is rank-deficient when the projection operator approaches zero as one approaches the equator and therefore its direct inverse is unstable.

To apply inverse theory, we adopt the Tikhonov formalism. The only aspect to note is that the data, model, and coefficient matrix are all complex. We introduce a model objective function as follows:

$$\varphi_m = \iint \left\{ \alpha_s |s(\rho)\widetilde{R}(u,v)|^2 + \alpha_u \left| \frac{\partial s(\rho)\widetilde{R}(u,v)}{\partial u} \right|^2 + \alpha_v \left| \frac{\partial s(\rho)\widetilde{R}(u,v)}{\partial v} \right|^2 \right\} du dv \quad [29]$$

where $s(\rho)$ is a spectral weighting function to ensure correct spectral shape of the recovered Fourier transform of the RTP field and $\rho = \sqrt{u^2 + v^2}$ is the radial wave number. Discretization based on the discrete Fourier transform yields the matrix form

$$\varphi_m = \vec{m}^{\,H} S^T W^T W S \vec{m} \quad [30]$$

where the superscript H denotes complex conjugate transposition, $W^T W$ is the finite difference representation of the three integrals in eqn [30] (see Li and Oldenburg, 1996), and S is a diagonal matrix containing the spectral weighting. The ideal spectral weighting would be the square root of the inverse of the radially averaged power spectrum of the data.

Similarly, we define a data misfit function

$$\varphi_d = \frac{1}{N} \sum_{i=1}^{N} \left| \widetilde{T}(u,v) - \widetilde{T}_i^{\text{pre}}(u,v) \right|^2 \quad [31]$$

where N is the number of discrete wave number points from the FFT and $\widetilde{T}_i^{\text{pre}}(u,v)$ is the predicted value.

Minimizing the total objective function $\varphi = \varphi_d + \beta\varphi_m$ using the calculus of variation, we obtain the matrix equation for the solution as

$$\left(G^H G + \beta S^T W^T W S\right) \vec{m} = G^H \vec{d} \quad [32]$$

Since $G^H G$ is a real-valued matrix, we have a system of equations consisting of a real matrix, with complex unknowns and a complex right-hand side. Any direct or iterative method can be used to obtain the final solution. The desired RTP solution is then obtained by an inverse Fourier transform.

We illustrate the approach with a synthetic example. Figure 11 shows the noisy total-field anomaly data at the equator and at the true RTP field produced by a rectangular prism with dimension of 10 by 10 by 1 unit buried at a depth of 1 unit. Using the methodology described earlier to carry out the RTP process, we obtain the RTP result shown in Figure 12, which clearly shows the central high above the causative body and four negative side lobes. There is no noticeable striping along the direction of declination, and the amplitude agrees well with the true value. Gradients in the east–west direction are weaker than their true values. The overall correspondence between the recovered RTP field and the true field at the pole shown in Figure 11 is excellent.

All methods mentioned earlier assume we know the direction of total magnetization. In the presence of a strong remanent component, the direction of magnetization is not known and one often uses a half reduction to the pole (HRTP) operator in which we filter the data using in eqn [25] only the term related to the known direction of measurement, that is,

$$\left(l\frac{iu}{\rho} + m\frac{iv}{\rho} + n \right) \quad [33]$$

to obtain the equivalent of having measured the vertical component of the magnetic field for the given magnetization, which is easier to interpret than total-field data.

All of these techniques assume that the directions of magnetization and ambient field are invariant over the entire survey area. While this is appropriate for many studies, it is not appropriate for large-scale studies, over which the Earth's magnetic field direction varies significantly, or in geologic environments, where remanent magnetization is important and variable. Arkani-Hamed (1988) addressed the former problem with an equivalent-layer scheme, in which variations in magnetization and ambient-field directions were treated as perturbations on uniform directions.

11.11.4.7 Pseudogravity Transformation

Baranov (1957) observed that one can use Poisson's relation

$$V = \frac{M}{Gd}\frac{\partial U}{\partial \alpha} = \frac{M}{Gd} g_\alpha \quad [34]$$

where V is the magnetic potential, U is the gravitational potential, d is the target density, g_α is the gravitational attraction in the direction of magnetization α, and G is the gravitational

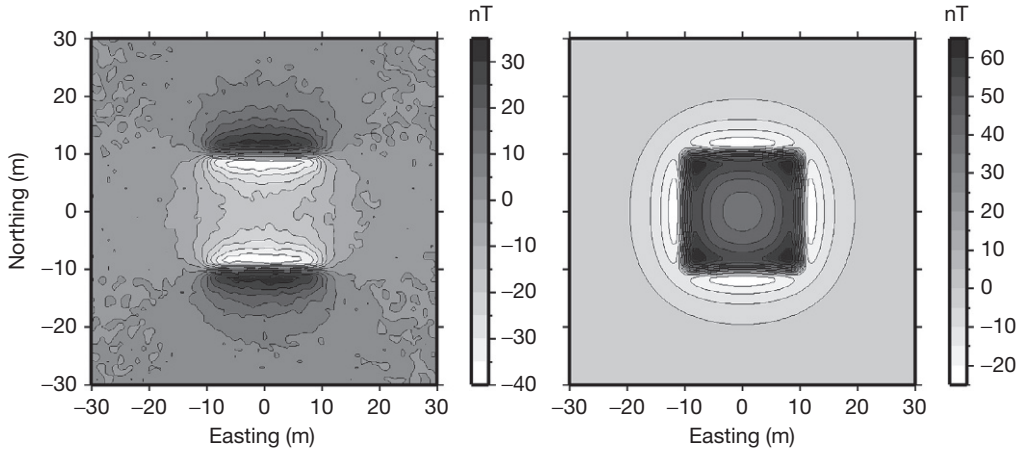


Figure 11 Total-field anomaly data at the equator (left) and at the true field at the pole (b). The contour interval is 5 nT. The causative body is a prism that has a width of 20 m in both horizontal directions, a thickness of 2 m, and a depth to the top of 1 m. The declination is 0° in both cases. The total-field anomaly at the equator has been contaminated by Gaussian random noise with zero mean and a standard deviation of 1 nT.

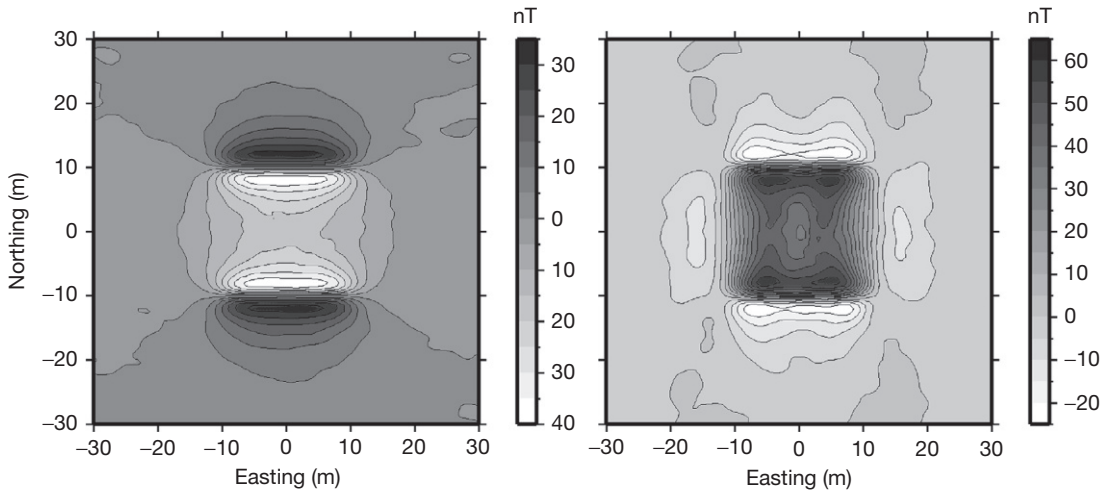


Figure 12 Result of RTP using the inverse formulation. The predicted data (left) reproduce the signal in the noisy total-field data in **Figure 11**. The recovered RTP data (right) are a good representation of the true RTP field shown in **Figure 11**. Reproduced from Li Y and Oldenburg DW (2001) Stable reduction to the pole at the magnetic equator. *Geophysics* 66: 571–578.

constant, to convert a magnetic anomaly to a corresponding gravity anomaly if we replace the magnetization distribution M with an equivalent density distribution d as long as we keep the ratio M/d constant. Since the resultant gravity anomaly has fictitious densities, Baranov named the output so obtained a pseudogravity map.

From the text mentioned earlier, we obtain, after taking the Fourier transform on both sides,

$$\widetilde{g}_{\alpha} = \frac{Gd}{M}\widetilde{V} \qquad [35]$$

To get the Fourier transform of g, we integrate the expression mentioned earlier along the direction of magnetization α to obtain

$$\widetilde{g} = \frac{Gd}{M}\frac{\widetilde{V}}{\rho\left(L\frac{iu}{\rho} + M\frac{iv}{\rho} + N\right)} \qquad [36]$$

Converting the magnetic potential V to total-field anomaly, we can write eqn [36] as

$$\widetilde{g} = \frac{Gd}{M}\frac{\widetilde{T}}{\rho\left(L\frac{iu}{\rho} + M\frac{iv}{\rho} + N\right)\cdot\left(l\frac{iu}{\rho} + m\frac{iv}{\rho} + nN\right)} \qquad [37]$$

This equation is similar to eqn [25] but has an additional term ρ in the denominator, which is equivalent to integration in the vertical direction. Thus, to obtain the pseudogravity field, we first reduce the data to the pole and then perform an integration in the vertical direction. All problems related to instability at low magnetic latitudes are still present. It is worth mentioning that the pseudogravity anomaly is identical to the magnetic potential.

The pseudogravity transformation has been commonly used as an interim step to several other edge-detection or depth-estimation techniques or in comparing with observed gravity anomalies.

11.11.4.8 Analytic Signal

For 2-D bodies of polygonal cross section, Nabighian (1972) showed that the horizontal and vertical derivatives of magnetic data are the Hilbert transform of each other. Thus,

$$\frac{\partial M}{\partial z} = H\left(\frac{\partial M}{\partial x}\right) \qquad [38]$$

where H signifies Hilbert transform. Taking the Fourier transform of eqn [38], we obtain

$$\frac{\partial \widetilde{M}}{\partial z} = -\mathrm{sgn}(\omega)\frac{\partial \widetilde{M}}{\partial x} \qquad [39]$$

where ω is the angular frequency and $\mathrm{sgn}(\omega)$ is the signum function.

Nabighian (1972) further defined a complex analytic signal of the form

$$A(x) = \frac{\partial M}{\partial x} + i\frac{\partial M}{\partial z} \qquad [40]$$

which, upon taking the Fourier transform, reduced to

$$\frac{\partial \widetilde{M}}{\partial x} + i\frac{\partial \widetilde{M}}{\partial z} = [1 + \mathrm{sgn}(\omega)]\frac{\partial \widetilde{M}}{\partial x} \qquad [41]$$

The amplitude of the analytic signal in eqn [40] that is defined by

$$a(x) = \sqrt{\left(\frac{\partial M}{\partial x}\right)^2 + \left(\frac{\partial M}{\partial z}\right)^2} \qquad [42]$$

yields a bell-shaped symmetrical function maximizing separately over each corner of the polygonal body. The half-maximum half-width of this bell-shaped function is equal to the depth to the respective corner (Nabighian, 1972). Note that the expression in eqn [42] is identical to the total gradient for the given data set. The amplitude of the analytic signal (the total gradient) has the remarkable property that it is independent of the magnetic inclination of the Earth's field or the presence of remanent magnetism. This property made it ideal

for interpreting, among others, the data collected over the Mid-Atlantic Ridge. The method works reasonably well as long as the various corners of the polygonal body are not too close to each other. If such is the case, then one can use a higher-order analytic signal amplitude, which is obtained from eqn [42] by replacing M with $\partial^n M/\partial x^n$, but this will make calculated data more sensitive to noise (Nabighian, 1972).

An example of using a higher-order analytic signal amplitude ($n=1$) is shown in **Figure 13** where the four bell-shaped functions are clearly visible over each corner of the trapezoid, with wider bell-shaped functions over the deeper corners.

For the 3-D case, Nabighian (1984) defined a complex analytic signal of the form

$$\frac{\partial \widetilde{M}}{\partial x}+\frac{\partial \widetilde{M}}{\partial y}+i\frac{\partial \widetilde{M}}{\partial z}=\left(1+\frac{u}{\sqrt{u^2+v^2}}\right)\frac{\partial \widetilde{M}}{\partial x}+\left(1+\frac{v}{\sqrt{u^2+v^2}}\right)\frac{\partial \widetilde{M}}{\partial y} \quad [43]$$

which is an extension of eqn [41] to 3-D. Note that for the 2-D case ($v=0$), eqn [43] reduces to eqn [41].

Unfortunately, none of the properties of the 2-D analytic signal transfer to the 3-D case. Roest et al. (1992) proposed to use again the total gradient

$$a(x)=\sqrt{\left(\frac{\partial M}{\partial x}\right)^2+\left(\frac{\partial M}{\partial y}\right)^2+\left(\frac{\partial M}{\partial z}\right)^2} \quad [44]$$

as the amplitude of the analytic signal. This amplitude however is strongly dependent on the inclination of the Earth's field and the presence of remanent magnetization, and only applying it to data that have already been reduced to the pole does minimize, but not eliminate, these effects (see also Li, 2006). The amplitude so defined is incorrectly known in the literature as the 3-D analytic signal. The analytic signal is a complex quantity as defined by eqn [43], while eqn [44] is a scalar quantity that defines the amplitude of the total gradient of the anomaly. As such, calculations based on eqn [44] represent the total gradient and not the amplitude of the analytic signal.

A slightly better expression than eqn [44] can be obtained by taking the square root of the sum of the squares of the four terms in eqn [43], which better positions the results over a target.

A thorough analysis of effects mentioned previously was given by Haney et al. (2003) and was based on the study of envelopes of 2-D and 3-D magnetic data. They showed that while in the 2-D case the total gradient is the envelope of both the horizontal and the vertical derivatives of the total-field anomaly over all possible directions of the Earth's field and source magnetization, this is not the case in 3-D. For the 3-D case, they found that the total gradient of the reduced-to-the-pole data defines the envelope only of the vertical derivative over all possible directions of the Earth's field and source magnetization, while the total gradient of the reduced-to-the-equator data defines the amplitude of only the horizontal derivative. This explains why the 2-D properties of the analytic signal cannot be extended to 3-D. Despite these shortcomings, using the total gradient, especially from reduced-to-the-pole data, can give diagnostic information on target boundaries, but caution must be exercised about also determining target depth.

A combined total gradient calculation using both eqns [42] and [44] over a shallow intrusive target of mining interest is shown in **Figure 14** where the ridgelike boundaries highlighted in red are located at a shallower depth and then the ridgelike boundaries highlighted in blue.

11.11.4.9 Amplitude Data

The amplitude of the analytic signal in the preceding section, which is equivalent to the total gradient, is based on the derivatives of the magnetic anomaly. An important property of the 2-D analytic signal is that its amplitude is the envelope of both its horizontal and vertical derivative functions

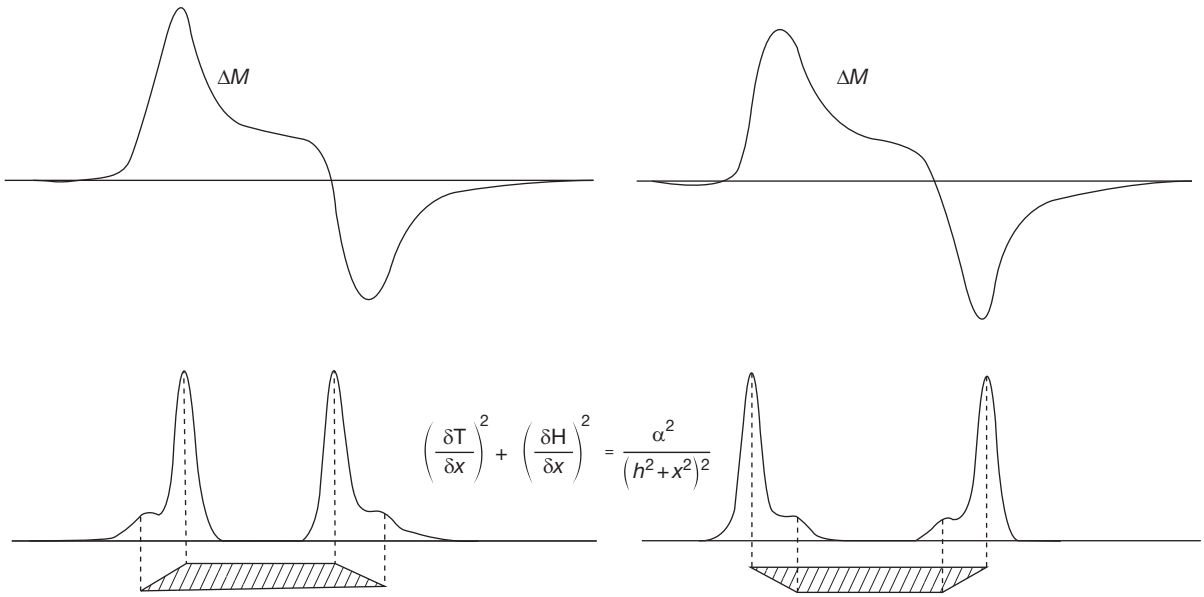


Figure 13 Total-field anomaly and analytic signal amplitude of order one over trapezoidal bodies. Reproduced from Nabighian MN (1974) Additional comments on the analytic signal of two-dimensional magnetic bodies with polygonal cross-section. *Geophysics* 39: 85–92.

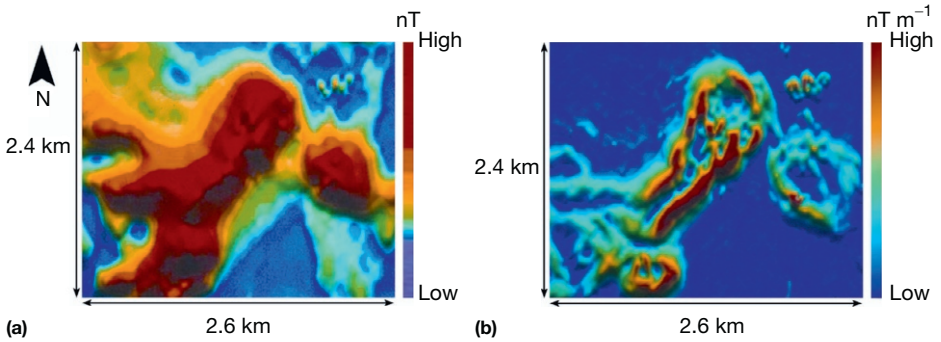


Figure 14 (a) Total-field magnetic map over a shallow intrusive target, (b) total gradient map.

(Haney et al., 2003). Although such a property does not extend to 3-D problems, the basic ideas of the analytic signal amplitude can be extended to the components of the magnetic anomaly itself. The amplitude of the anomalous magnetic field vector, $B_a = |\mathbf{B}_a|$, is independent on the magnetization direction in 2-D problems (Nabighian, 1972), while in 3-D problems, the amplitude data are only weakly dependent on magnetization direction. For this reason, Stavrev and Gerovska (2000) introduced amplitude data as one of the several direction invariant transformations of magnetic data in tackling the difficulties associated with an unknown magnetization direction. This property of amplitude data enables their use in the presence of strong remanent magnetization and provides the opportunity for direct mapping of source locations and for 3-D inversion for the magnetization magnitude without knowing its direction (Shearer, 2005). We illustrate the nearly invariant properties of the amplitude data here and defer the discussion on inversion to **Section 11.11.8**.

Figure 15 illustrates the weak dependence of amplitude data on the magnetization direction. The causative body consists of two segments with different strike directions and variable depth to the top. The inducing field has an inclination of $I = -5^\circ$ and a declination of $D = 0^\circ$. The column on the left shows the total-field anomaly for two different magnetization directions, one aligned with the inducing field and another nearly perpendicular. The second case emulates the presence of strong remanent magnetization. The column on the right shows the corresponding amplitude data. While the total-field anomalies in the two cases are noticeably different, the amplitude data are much more similar and provide a good indication of the horizontal location of the causative body. Furthermore, in the case of remanent magnetization, the RTP process may not be applicable in practice because of the unknown magnetization direction.

The amplitude data preserve the low wave number content in the data and therefore better retain the signal emanating from deeper causative bodies. Consequently, the amplitude data may be a better alternative than gradient data when deeper targets are of interest.

11.11.4.10 Source Parameter Imaging Method

Thurston and Smith (1997) extended the theory of the complex analytic signal given in eqn [40] by computing three complex parameters (local amplitude, local phase, and local frequency) from which they can then calculate various source parameters.

They started by writing eqn [40] in the equivalent form

$$A(x) = \frac{\partial M}{\partial x} + i\frac{\partial M}{\partial z} = |A|e^{i\theta} \qquad [45]$$

where θ is the local phase

$$\theta = \tan^{-1}\left(\frac{\partial M}{\partial z} \Big/ \frac{\partial M}{\partial x}\right) \qquad [46]$$

and the local amplitude is

$$|A| = \sqrt{\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2} \qquad [47]$$

Further, they introduced the local frequency f that they defined as the rate of change of local phase θ with respect to x:

$$f = \frac{1}{2\pi}\frac{\partial}{\partial x}\tan^{-1}\left(\frac{\partial M}{\partial z} \Big/ \frac{\partial M}{\partial x}\right) \qquad [48]$$

Carrying out the differentiation in eqn [48] and after converting the local frequency to local wave number k_1 ($k_1 = 2\pi f$), they obtained

$$k_1 = \frac{1}{|A|^2}\left(\frac{\partial^2 M}{\partial x \partial z}\frac{\partial M}{\partial x} - \frac{\partial^2 M}{\partial x^2}\frac{\partial M}{\partial z}\right) \qquad [49]$$

Smith et al. (1998) extended the results mentioned previously by defining a second-order analytic signal defined by

$$A_2(x) = \frac{\partial^2 M}{\partial x \partial z} + i\frac{\partial^2 M}{\partial z^2} \qquad [50]$$

and a second-order local wave number

$$k_2 = \frac{\partial}{\partial x}\tan^{-1}\left(\frac{\partial^2 M}{\partial z^2} \Big/ \frac{\partial^2 M}{\partial x \partial z}\right) \qquad [51]$$

Using the definition mentioned earlier for k_1 and k_2, Smith et al. (1998) related them to the target depth via the structural index (SI) N (see **Section 11.11.4.16**):

$$k_1 = \frac{(N+1)h}{(x-x_0)^2 + h^2}$$
$$k_2 = \frac{(N+2)h}{(x-x_0)^2 + h^2} \qquad [52]$$

where x_0 is the horizontal location of the target. The depth h is dependent on target type via the SI. The value of the SI N is

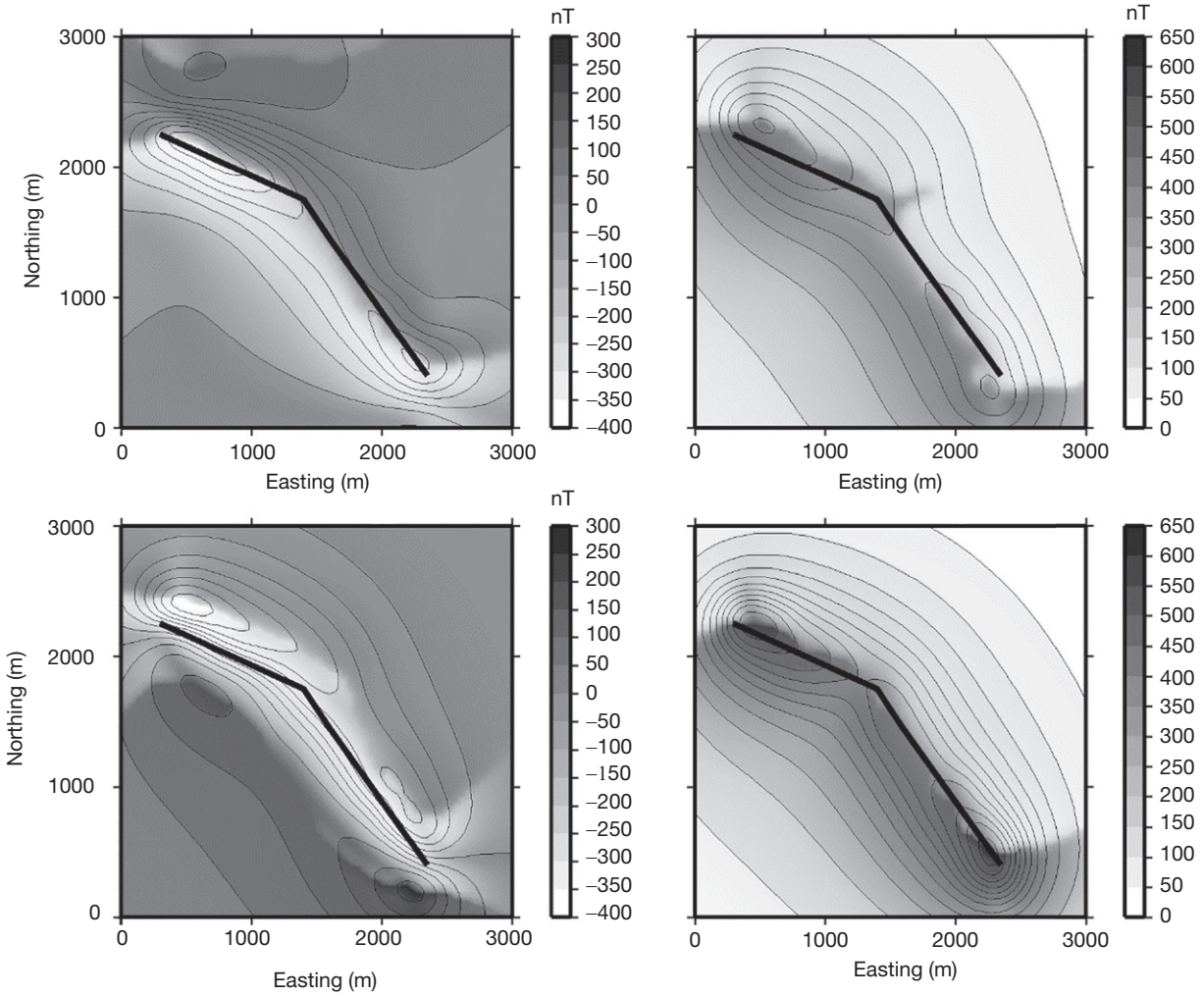


Figure 15 A synthetic example illustrating the weak dependence of amplitude data on magnetization direction and their utility as a mapping tool in place of the traditional RTP data. The location of the causative body is indicated by the solid line segments. The inducing field has an inclination of $I = -5^\circ$ and a declination of $D = 0^\circ$. The top left panel has a magnetization direction aligned with the inducing field, while the lower left panel has a magnetization direction with an inclination of $I_m = 75^\circ$ and a declination $D_m = 45^\circ$, which could be a scenario with strong remanent magnetization. While the RTP process may not be applicable in the latter case because of the unknown magnetization, in practice, the amplitude data provide a good indication of the horizontal location of the causative body.

0 for the top of a contact or fault with large depth extent, one for the top edge of a thin sheet, two for the axis of a horizontal cylinder, and three for the center of a sphere. In addition to depth determinations, Thurston and Smith (1997) also showed how to determine the target dip and susceptibility for sloping contacts or thin dikes.

The source parameter imaging method, also known as the local wave number method, gives reliable depth determination for simple sources but fails when various target corners are close to each other (e.g., a thick dike). In addition, due to the use of higher-order derivatives, the method is sensitive both to noise and to aliasing effects.

11.11.4.11 Tilt Angle Method

For 2-D profiles, Miller and Singh (1994) defined the tilt angle as the ratio between the vertical and the total horizontal derivatives:

$$\theta = \tan^{-1} \frac{\dfrac{\partial M}{\partial z}}{\dfrac{\partial M}{\partial h}} \qquad [53]$$

where

$$\frac{\partial M}{\partial h} = \sqrt{\left(\frac{\partial M}{\partial x}\right)^2 + \left(\frac{\partial M}{\partial y}\right)^2} \qquad [54]$$

The tilt angle varies between $-\pi/2$ and $\pi/2$. The authors stated that using the ratio between two derivatives acts like an automatic gain-control filter and deeper edges can thus be more easily identified. The method was further extended by Salem et al. (2008) by using the derivatives of the tilt angle and obtaining simple equations similar to the Euler deconvolution equations (see succeeding text).

11.11.4.12 Total Horizontal Gradient

Cordell (1979) proposed the THG method, initially for gravity data, based on the observation that a near-vertical fault produces a gravity anomaly whose horizontal gradient is largest almost directly over the top edge of the fault. As a result, the horizontal gradient will plot as a ridge whose maxima are located approximately over target edges. The method is applicable to magnetic data also if one first reduces the data to the pole or better if one calculates the pseudogravity anomaly. The THG is calculated using the expression

$$\mathrm{THG} = \sqrt{\left(\frac{\partial M}{\partial x}\right)^2 + \left(\frac{\partial M}{\partial y}\right)^2} \qquad [55]$$

Blakely and Simpson (1986) presented a useful method for automatically locating and characterizing the crests of the horizontal gradient magnitude. An example of the THG over the same shallow intrusive target as in Figure 15 is shown in Figure 16. When compared with Figure 15(b), one can see that the THG map, while determining quite accurately the target boundaries, does not give any information about the relative target depths in various areas of the target.

A different approach to edge detection was published by Cooper and Cowan (2008) by using normalized standard deviation, which is based on ratios of the windowed standard deviation of derivatives of the field.

11.11.4.13 Terracing

With the aim of producing a map composed of uniform domains separated by abrupt boundaries, Cordell and McCafferty (1989) developed the terracing operator, which is an iterative process based on the measured field's local curvature. The terracing operator is normally applied to either the reduced-to-the-pole or the pseudogravity data.

In this technique, a three-point window is moved across the data, and the value of the field at the center of the window is increased, decreased, or unchanged, depending on the algebraic sign of local curvature of the function evaluated at the central point. The local curvature at station i is given by the second derivative and is calculated using finite differences:

$$C_i = \frac{M_{i+1} - 2M_i + M_{i-1}}{\Delta x^2} \qquad [56]$$

where Δx is the sampling interval. If the curvature is positive, the central point is assigned the lower value of the two endpoints, and if the curvature is negative, the central point is assigned the higher value of the two endpoints (see Figure 17). When the curvature is zero, the central point is left unchanged.

The three-point window is passed point by point over the whole profile, and this process is repeated iteratively until the output contains only horizontal and vertical segments. The procedure is identical for map data where the window is

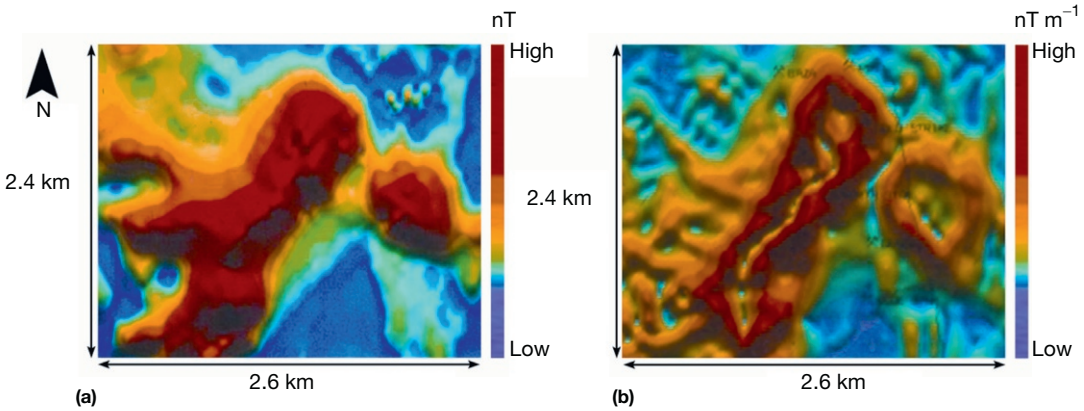


Figure 16 (a) Total-field map over an intrusive target shown in Figure 14, (b) total horizontal gradient map.

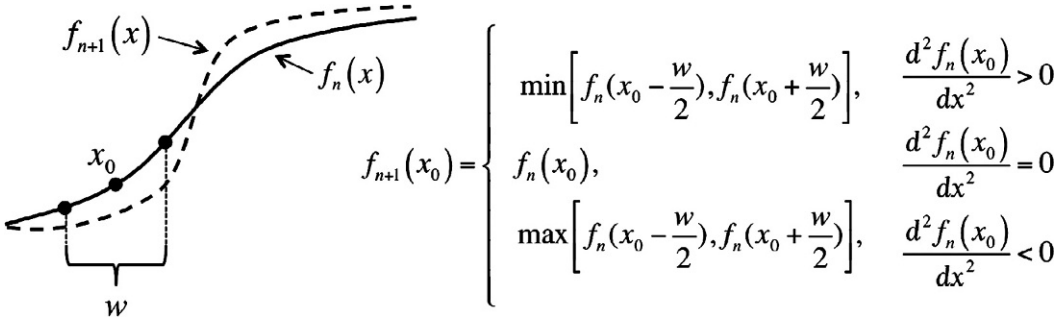

$f_{n+1}(x)$
$f_n(x)$
x_0
w
$f_{n+1}(x_0) = \begin{cases} \min\left[f_n(x_0 - \frac{w}{2}), f_n(x_0 + \frac{w}{2})\right], & \frac{d^2 f_n(x_0)}{dx^2} > 0 \\ f_n(x_0), & \frac{d^2 f_n(x_0)}{dx^2} = 0 \\ \max\left[f_n(x_0 - \frac{w}{2}), f_n(x_0 + \frac{w}{2})\right], & \frac{d^2 f_n(x_0)}{dx^2} < 0 \end{cases}$

Figure 17 Terracing operator. The midpoint value of a three-point window is increased or decreased depending on whether the curvature is negative or positive.

now of size 3×3 and the curvature is determined by fitting a second-order polynomial through the nine points and calculating the curvature analytically.

The terraced function simulates zones of different susceptibilities, and when various adjacent terraced functions are combined, they evoke images of rice paddies, hence the name.

The terracing operator only outlines zones with different magnetic susceptibilities without assigning a susceptibility to each terraced area. To partially overcome this, an unknown susceptibility is assigned to each terraced area, and using a linear approach, one determines the best susceptibility value for each terraced area such that the calculated output matches as closely as possible the measured data, thus yielding a terraced susceptibility map.

An example of a terraced susceptibility map over a large area flown at low magnetic latitudes is shown in Figure 18.

11.11.4.14 Matched Filtering

In potential-field methods, matched filtering has been primarily used to separate data into anomaly components representing different source depths. The method was developed for use with magnetic data by Spector (1968) and Spector and Grant (1970) who showed that a statistical ensemble of sources at a depth z below the Earth's surface results in a Fourier amplitude spectrum on the surface of the Earth given by

$$\widetilde{M}(u, v) = A|\rho|^n e^{-\rho z} \qquad [57]$$

where A is a constant proportional to the amplitude of the white-noise magnetization, $n = 1$ for an ensemble of finite thickness (a magnetic dipole layer), and $n = 0$ for an ensemble of infinite depth extent (magnetic half-space). From eqn [57], the logarithmic radial power spectra of aeromagnetic data will contain straight slope segments that can be interpreted as arising from statistical ensembles of sources at different depths. Spector (1968) applied matched filtering in both frequency and space domains. If the spectrum of the expected signal is known, matched filtering can help locate the signal in a given data set since the matched filter has the same spectrum as the desired signal. Syberg (1972) introduced the term *matched filter* and applied the technique to modeling azimuthal variations within each band pass. Ridsdill-Smith (1998a,b) developed wavelet-based matched filters, while Phillips (2001) generalized the Fourier approach of Syberg (1972) to sources at more than two depths and explained how matched Wiener filters could be used as an alternative to the more common amplitude filters.

An alternative to matched filters was developed by Jacobsen (1987) based on differencing upward-continued fields. Cowan and Cowan (1993) compared the results of Spector's matched filter, the Cordell filter, the Jacobsen filter, and a second vertical derivative filter on an aeromagnetic data set from Western Australia.

11.11.4.15 Werner Deconvolution

The Werner deconvolution (Werner, 1955) interprets the location and depth of magnetic contacts or dikes. For 2-D targets, depths are calculated independent of target strike, dip, and remanent magnetization; and the calculations are valid at any magnetic inclination.

For a thin dike, the magnetic anomaly can be written as (Nabighian, 1972)

$$M(x) = \alpha \frac{h\cos\phi + (x - x_0)\sin\phi}{h^2 + (x - x_0)^2} \qquad [58]$$

where x_0 is the horizontal location of the top of the dike, h is the depth to the top of the dike, φ depends on the dike's dip and the total magnetization (induction plus remanent) of the Earth's field, and α depends on the Earth's magnetic field intensity, the dike's dip, the susceptibility contrast of the dike, and the survey profile orientation. By eliminating the denominator, the expression mentioned earlier becomes

$$M(x)\left[h^2 + (x - x_0)^2\right] = \alpha[h\cos\phi + (x - x_0)\sin\phi] \qquad [59]$$

The expression mentioned previously has four unknowns: x_0, h, α, and φ. Applying eqn [59] to four consecutive points (x_1, x_2, x_3, x_4) on a profile, we obtain four equations with four

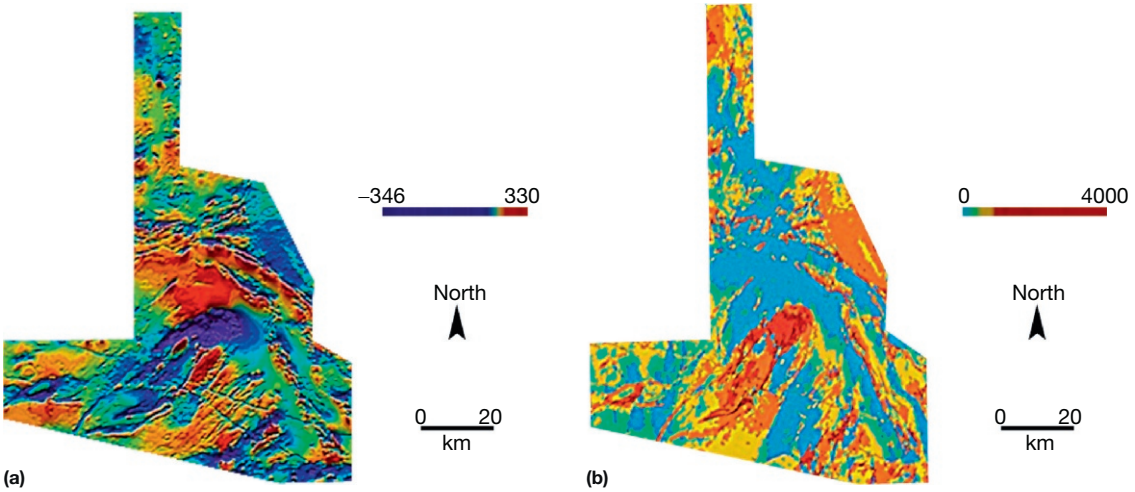


Figure 18 (a) Total-field map over an area at low magnetic latitude, (b) terraced susceptibility.

unknowns that can be solved for the four unknown parameters. The method is repeated by moving the four points over by one station spacing (x_2, x_3, x_4, x_5). If the data from all these points came from the same anomaly, we should get an identical result. By carrying out this procedure over the whole profile, we should get clustering over the top of each thin dike present along the profile.

To account for the fact that there could be overlapping anomalies, eqn [58] is modified by adding a parabola to the thin dike anomaly to approximate the effect of the neighboring anomaly:

$$M(x) = \alpha \frac{h\cos\phi + (x - x_0)\sin\phi}{h^2 + (x - x_0)^2} + A + Bx + Cx^2 \qquad [60]$$

Clearing the denominator as before, we now have seven unknowns that can be determined by using this equation over seven consecutive points. This approach was implemented by Hartman et al. (1971). The addition of a parabola to account for the neighboring anomaly did help improve results, but the output still did not exhibit clear clusters and 'wormlike' features were prominently displayed. The method has been used with both total-field data and its horizontal derivative.

Hansen and Simmonds (1993) made a major contribution in this area by developing the theory of multisource Werner deconvolution that accounts simultaneously for the presence of many sources without the need for a parabolic term. Their approach starts with the definition of the complex analytic signal given by eqn [40] and notes that for a thin magnetic dike whose vertex is at the point (x_1, z_1), the analytic signal can be written as

$$A_1(x) = \frac{\alpha_1}{x - w_1} \qquad [61]$$

where $w_1 = x_1 + iz_1$ and α_1 is a complex constant as defined previously. They further noted that for a thin dike, constants a_0 and b_0 can be found to satisfy the relation

$$a_0 + b_0 A_1(x) + xA_1(x) = 0 \qquad [62]$$

After clearing the denominator in eqn [62], we can write

$$xA_1(x) - x_1 A_1(x) - izA_1(x) = \alpha \qquad [63]$$

By comparing eqn [62] with eqn [61], it follows that

$$a_0 = -\alpha_1 \quad \text{and} \quad b_0 = -\omega_1 \qquad [64]$$

Thus, for a single source, using the analytic signal, one determines first the constants a_0 and b_0 from which we then determine w_1 and α_1.

For the general case of n sources, eqn [61] can be written as

$$A_n(x) = \sum_{i=1}^{n} \frac{\alpha_i}{x - w_i} \qquad [65]$$

and, after a lengthy derivation, the extension of eqn [63] becomes

$$\sum_{i=0}^{n-1} a_i x^i + \sum_{i=0}^{n-1} b_i x^i A_n(x) + x^n A_n(x) = 0 \qquad [66]$$

Finally, eqn [65] can be extended for n sources with vertices at α_j and w_j leading to

$$\sum_{i=0}^{n-1} b_i w_j^i + w_j^n = 0 \quad \text{and} \quad \alpha_j = -\frac{\sum_{i=0}^{n-1} a_i w_j^i}{\prod_{i=1}^{n} (w_j - w_i)} \quad \text{for} \quad i \neq j \qquad [67]$$

The first equation mentioned earlier has n roots for w that determine later the n roots for α. Figure 19 shows the theoretical results obtained over two trapezoidal bodies using both the single-source Werner deconvolution technique and the multisource Werner deconvolution technique. As expected, the multisource technique yielded only clusters concentrated over each corner of the target, whereas the single-source technique properly identified only the shallower corners.

11.11.4.16 Euler Deconvolution

The Euler deconvolution has been widely used in automatic magnetic interpretations because it requires no prior knowledge of source magnetization direction. Despite the fact that the Euler deconvolution is only valid for homogenous functions and only very few functions used in magnetic exploration are homogenous, geologically plausible results can still be obtained in many cases with real data.

A function $f\left(\vec{x}\right)$ of a set of variables $\vec{x} = (x_1, x_2, x_3 \ldots.)$ is homogenous if it satisfies

$$f\left(t\,\vec{x}\right) = t^n f\left(\vec{x}\right) \qquad [68]$$

where t is an arbitrary constant and n is the degree of homogeneity. (A function of the type $f = A/r^n$ is homogeneous of degree $-n$. To avoid negative numbers for the degree of homogeneity, we define a positive SI as $N = -n$.) A function that is homogenous satisfies the equation (Hood, 1965)

$$\vec{x} \cdot \nabla f\left(\vec{x}\right) = nf\left(\vec{x}\right) \qquad [69]$$

Hood (1965) was the first to write down the Euler homogeneity equation for the magnetic case, and he derived the SI for a point pole and for a point dipole. The method was further developed by Thompson (1982) and later on applied to 3-D problems by Reid et al. (1990).

To determine the location and depth to a given target using eqn [69], one needs to know the three orthogonal gradients of a given magnetic component as well as the magnetic component itself. As an example, for a given magnetic component M, the standard Euler deconvolution equation can be written as

$$(x - x_0)M_x + (y - y_0)M_y + (z - z_0)M_z = -N(B - M) \qquad [70]$$

where x_0, y_0, and z_0 are the unknown coordinates of the source body center or its top edge; x, y, and z are the known coordinates of the observation point; B represents an unknown constant background (regional) value for the component M; and N is the SI whose value depends on source type.

For some simple targets, the SI values are shown in Table 2.

Only for the four bodies shown in Table 2 is the Euler equation homogeneous. Other possible sources are not strictly homogeneous, and the application of the Euler deconvolution to such cases is at best an approximation.

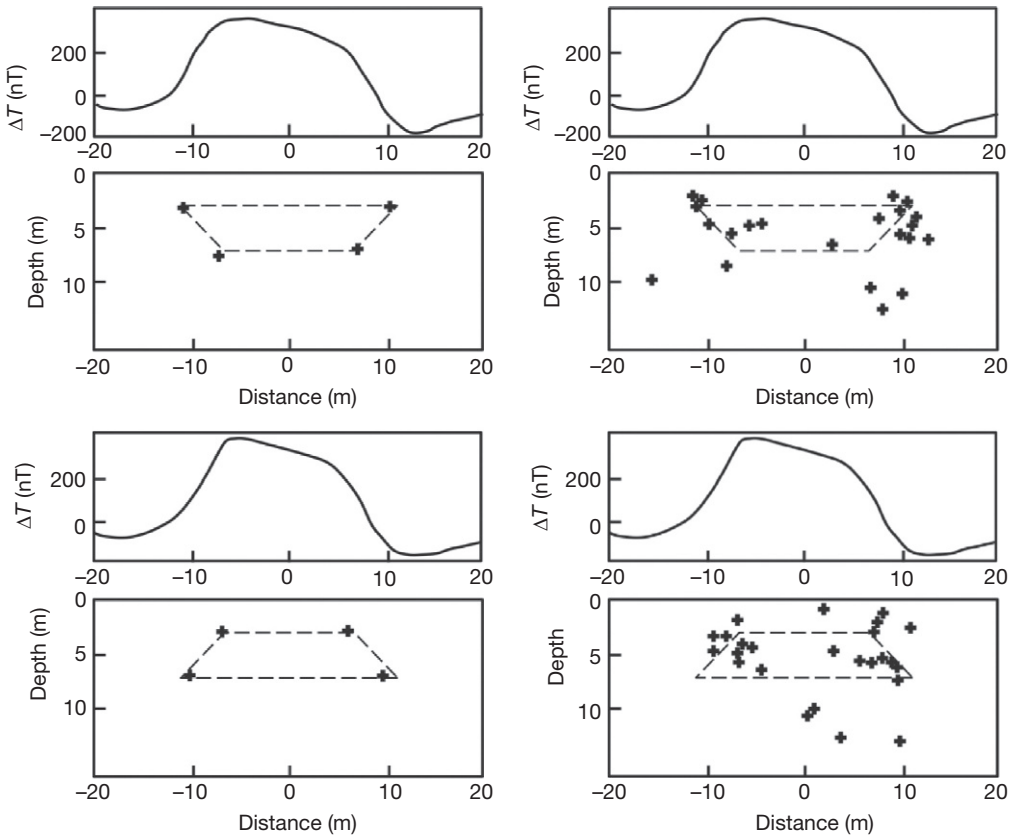


Figure 19 Theoretical calculations over two trapezoidal bodies using the standard, single-source (left) and the extended, multisource (right) Werner deconvolution.

Table 2 Structural indexes for various targets

Body type	*Structural index (SI)*
Sphere	3
Horizontal cylinder	2
Thin dike	1
Contact	0

The technique is normally applied to gridded data. By applying the Euler equation to every point on an $n \times n$ window, we obtain an overdetermined system of equations from which the unknown quantities mentioned previously can be determined. For the standard Euler application, we have to assume an SI N (in the standard Euler deconvolution, the target depth z_0 and the SI N of the source cannot be estimated simultaneously because they are linearly dependent) and then solve for the remaining four unknown parameters x_0, y_0, z_0, and B. In the absence of other supporting data about the target shape, an SI N of 0.5 is usually applied for the most general case.

The process is repeated for successive window positions, and the solutions are then clustered by both horizontal location and depth for a cleaner representation.

For the 2-D case, Mushayandebvu et al. (2001) introduced an additional equation in the formulation mentioned previously that expressed the transformation of homogeneous functions under rotation, and they showed that the combined two equations led to better source parameter estimation. Mushayandebvu et al. (2001) named this approach the extended Euler deconvolution. Shortly afterward, Nabighian and Hansen (2001) showed that the standard Euler eqn [70] is also satisfied in 3-D if we replace the field component M with the X- and Y-components of its Hilbert transform (Nabighian, 1984):

$$\begin{aligned}(x-x_0)\mathbf{H}_x + (y-y_0)\mathbf{H}_x + (z-z_0)\mathbf{H}_x &= -N\mathbf{H}_x \\ (x-x_0)\mathbf{H}_y + (y-y_0)\mathbf{H}_y + (z-z_0)\mathbf{H}_y &= -N\mathbf{H}_y\end{aligned} \quad [71]$$

This added two more equations to the standard Euler equation that led to more stable solutions and also allowed for the independent determination of the SI N since the Hilbert transform of a constant background is zero. Thus, in using the extended Euler deconvolution, we can consider the SI N also as an unknown parameter and we solve now for five unknowns x_0, y_0, z_0, B, and N.

The extension to gravity tensor gradient data was published by Zhang et al. (2000), which could be extended to magnetic data. The tensor Euler deconvolution uses all magnetic gradient tensor components and all components of the magnetic anomaly vector. In addition to the standard equation mentioned earlier, it also uses two similar equations for the horizontal components:

$$(x-x_0)U_{xx}+(y-y_0)U_{xy}+(z-z_0)U_{xz}=-N(B_x-U_x)$$
$$(x-x_0)U_{yx}+(y-y_0)U_{yy}+(z-z_0)U_{yz}=-N(B_y-U_y) \quad [72]$$

where, for example,

$$U_{xy}=\frac{\partial^2 U}{\partial x\partial y} \quad [73]$$

The solution process is similar to the one mentioned earlier with the exception that we now have three times more equations than before, similar to the case for the extended Euler deconvolution. Assuming a value for the SI N, we now have to solve for the six unknowns parameters x_0, y_0, z_0, B_x, B_y, and B_z. The tensor Euler solutions are generally more tightly clustered and define linear features better than their conventional counterparts, but they do require inputting an estimated value for the SI N.

Hansen and Suciu (2002) extended the standard Euler deconvolution to multiple sources, and the reader is referred to this chapter for further details. The method does require that all sources in a given window have the same SI and its implementation requires a very careful determination of the coefficients of a linear equation. Although very promising, the multisource Euler deconvolution is not presently part of the standard arsenal of a geophysicist.

An interesting case history was published by Reid et al. (1990) over an area in the south-central part of Great Britain, the results of which were summarized by Reynolds (1997). The Euler solutions shown in 20 have outlined the two north–south boundary faults marking the central block, and the depth to the Eastern Boundary fault was determined within 0.3 km compared with the depths determined by other means (**Figure 20**).

On the other end of the scale, the extended Euler deconvolution (Nabighian and Hansen, 2001) and its ability to calculate the structure index have been used successfully in detecting buried UXOs (**Figure 21**). The premise is that UXO targets are confined metallic objects and behave like dipole sources. Therefore, the corresponding magnetic field should have an SI close to three (Davis et al., 2010). **Figure 21** illustrates this approach. **Figure 21(a)** displays a set of total-field magnetic data in grayscale contours in the background and the locations of UXO targets in white circles, and detection results in plus (+) as overlays. Using a threshold of 2.7 to measure the closeness of the calculated SI to the theoretical value of 3.0, the algorithm detects 19 out of the 20 targets. **Figure 21(b)** displays the calculated SI for the 20 dipolar anomalies.

11.11.4.17 Wavelet Analyses

The wavelet transform is emerging as an important processing technique in potential-field methods and has contributed significantly to the processing and inversion of magnetic data. The primary reason is the localization property of the wavelet transform. The concept of the continuous-wavelet transform (CWT) was introduced initially in seismic data processing (Goupillaud et al., 1984), while a form of discrete wavelet transform has long been used in communication theory. These were unified through an explosion of theoretical developments in applied mathematics. Potential-field analyses – magnetic methods in particular – have benefited greatly from these developments.

There are three major categories of wavelet applications in magnetic methods. The first is the use of natural wavelets derived from Green's function of the magnetic field to extract source parameters. Moreau et al. (1997) were the first to use the CWT to analyze potential-field data for simple monopole sources. In a seminal paper, Hornby et al. (1999) independently developed a similar theory and recast commonly used processing methods in potential fields in terms of CWT. These wavelets are essentially different second-order derivatives of the potential produced by a point monopole source. The translation of the wavelet corresponding to shifting the monopole horizontally and dilation corresponding to choosing different 'observation' heights yield the wavelet family. Methods based on the CWT identify locations and boundaries of causative bodies by tracking the extrema of wavelet transforms. Martelet et al. (2001) applied a similar wavelet transform to magnetic data to identify geologic boundaries. Haney and Li (2002) used the CWT to estimate total magnetization directions in the presence of remanent magnetization.

A second class of methods is applied primarily for data processing that uses discrete wavelet transforms based on compactly supported orthonormal wavelets (Daubechies, 1992). Chapin (1997) applied the wavelet transforms to the interpretation of magnetic profiles. Ridsdill-Smith and Dentith (1999) used wavelet transforms to enhance the processing of aeromagnetic data. Lyrio et al. (2004) improved upon the concept of wavelet denoising in signal processing and applied it to processing of the gravity gradiometry data by first estimating the noise level in the wavelet domain and then removing the noise accordingly. The same approach has been applied to the denoising of magnetic data. This approach has also been used to carry out magnetic anomaly separation by Krahenbuhl et al. (2011) to extract anomalies due to buried metallic targets from the background responses due to magnetic units in the host geology.

The third category is to use discrete wavelet transforms to compress the operator and improve the numerical efficiency in equivalent source construction and inversion of magnetic data. This approach has also been applied to 3-D generalized inversion. This class will be discussed in more detail later.

11.11.5 Magnetic Forward Modeling

For quantitative interpretation of magnetic data, it is necessary to be able to carry out forward modeling of a given magnetic source distribution in the subsurface. A great deal has been published on this subject. Grant and West (1965) and Telford et al. (1996) had an excellent collection of analytic solutions for geometrically simple objects with uniform magnetization within. Modeling methods based on magnetic anomalies across 2-D bodies of polygonal cross section (Talwani and Heirtzler, 1964) and over 3-D bodies represented by right rectangular prisms (Bhattacharyya, 1964; Bott, 1963), by polygonal faces (Bott, 1963), or by stacked thin horizontal sheets of polygonal shape (Talwani, 1965) have played important role in quantitative interpretation. The 2-D magnetic forward modeling algorithm of Talwani and Heirtzler (1964) was later modified to

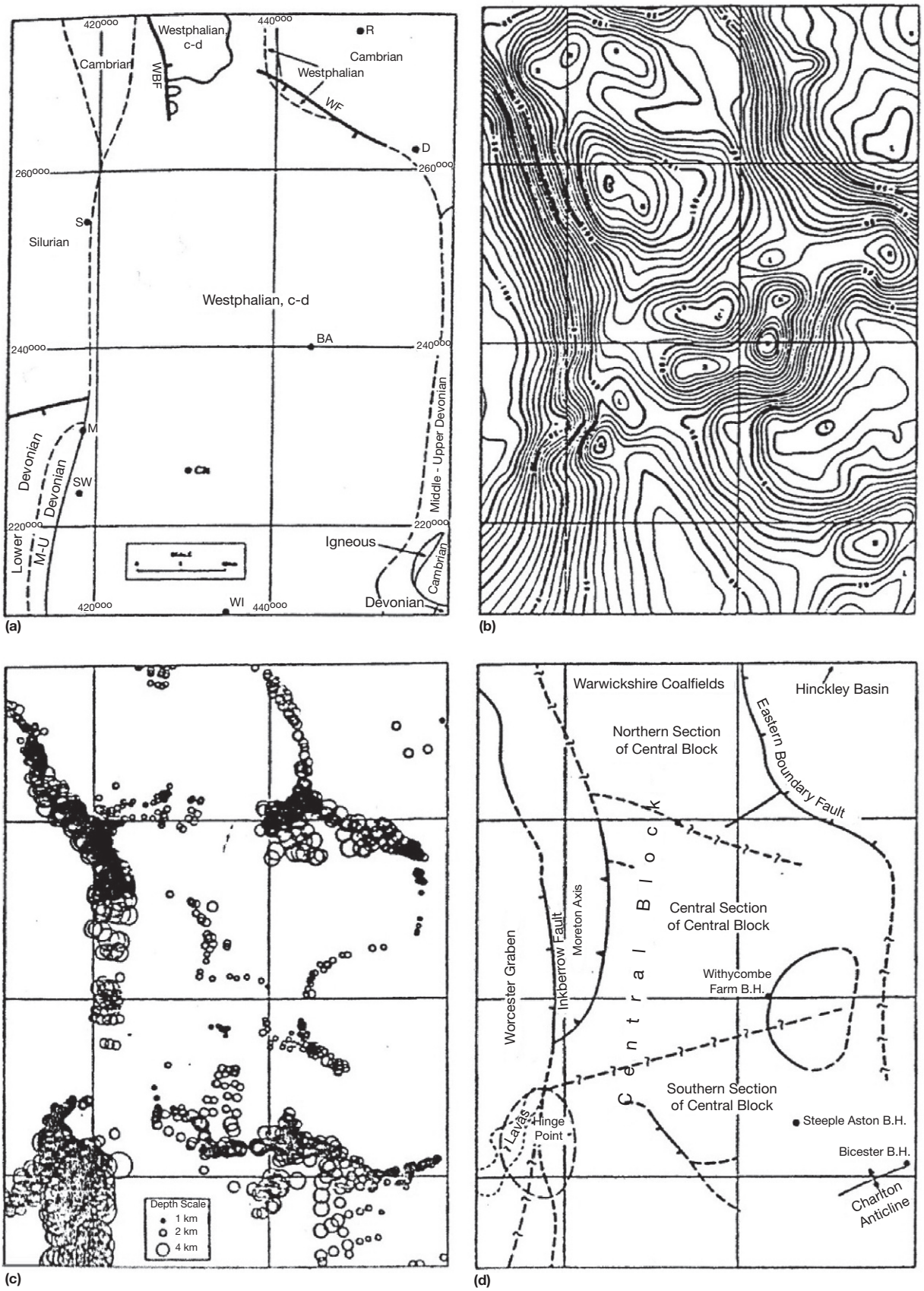


Figure 20 (a) Simplified geologic map of southeastern part of Great Britain. (b) Total-field aeromagnetic map over the area. (c) The Euler solutions derived using a structural index of one. (d) Structural interpretation of the Euler solutions. Reproduced from Reid AB, Allsop JM, Granser H, Millett AJ, and Somerton IW (1990) Magnetic interpretation in three dimensions using Euler deconvolution. *Geophysics* 55: 80–91; Reynolds JM (1997) An Introduction to Applied and Environmental Geophysics. John Wiley & Sons.

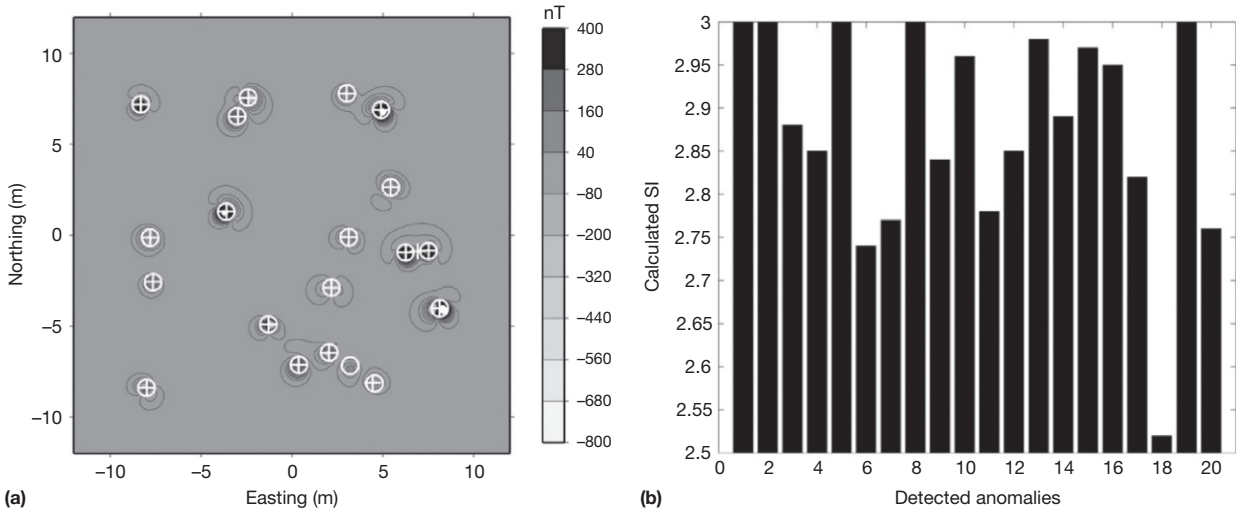


Figure 21 A synthetic example of automatic UXO detection using the extended Euler deconvolution. (a) Synthetic magnetic data in grayscale contours and the corresponding UXO target (white circles) and the detection results (+). (b) The corresponding UXO detection result showing the computed structural indexes for the 20 anomalies. Any anomaly with a structural index >2.7 (the cutoff value) was considered indicative of a UXO. The method identifies 19 out the 20 targets. Reproduced from Davis K, Li Y, and Nabighian M (2010) Automatic detection of UXO magnetic anomalies using extended Euler deconvolution. *Geophysics* 75 (3), G13–G20.

include bodies of finite strike length (Cady, 1980; Rasmussen and Pedersen, 1979; Shuey and Pasquale, 1973). Barnett (1976) used the approach of Bott (1963) to modeling 3-D bodies by using triangular facets and by Okabe (1979).

Because of the linear superposition of magnetic fields, Fourier transforms of magnetic fields of simple sources (Bhattacharyya, 1966; Blakely, 1995; Spector and Bhattacharyya, 1966; Pedersen, 1978) provide an elegant and efficient means for forward modeling. Parker (1972) presented a practical Fourier method for modeling complexly layered gravity and magnetic sources, where the observations are on a flat plane or other surface that is above all the sources.

The venerable right rectangular prism has remained popular for voxel-based magnetic forward modeling and inversion. Many widely used forward modeling and inversion algorithms have made extensive use of this simple source geometry by discretizing the subsurface volume into a set of contiguous rectangular prisms using a 3-D orthogonal mesh (e.g., Bhattacharyya, 1980; Guillen and Menichetti, 1983; Li and Oldenburg, 1996; Pilkington, 1997; Zeyen and Pous, 1991). The total response of a model is obtained by the summation of the responses due to all prisms. The solution for a single prism has been presented extensively in the literature (Bhattacharyya, 1964; Rao and Babu, 1991; Sharma, 1966). This approach has a strong connection with the Fourier-domain approach based on the prismatic cells and convolution theorem but allows the flexibility of variable discretization intervals and arbitrary observation locations at different elevations.

Parallel with the development of numerical simulation techniques in geophysical methods, techniques based on the numerical solution of differential equations have been applied to the forward modeling of magnetic data. Lelièvre (2005) and Lelievre and Oldenburg (2006) used the approach to model the magnetic response in the presence of high magnetic susceptibility. Davis et al. (2013) used this approach to carry out forward modeling in the inversion of regional-scale magnetic data.

Among these different approaches, the use of rectangular prisms with a given magnetization (see **Section 11.11.2**) still plays an important role and is widely used in generalized inversions. For this reason, we discuss several aspects of its practical utility in the succeeding text. Readers are referred to the literature for descriptions on other approaches.

Hjelt (1972) presented the equations for the magnetic field of a dipping prism, having two opposite vertical sides that are parallelograms. This particular form of voxel is useful for modeling magnetic anomalies that have resulted from distortions of layered media by geologic processes such as faulting and folding (Jessell, 2001; Jessell and Valenta, 1996).

11.11.5.1 Magnetic Anomaly

Although the magnetic field is not strictly a conservative field, it is curl-free in the source-free regions. Consequently, we can define the magnetic field, specifically the magnetic flux density, as the negative gradient of a scalar potential. This approach allows us to describe the magnetic data by Laplace's equation just as in the case of gravity data. Equally importantly, the introduction of a scalar potential allows us to treat a small current loop as a magnetic dipole defined by the product of the current and the area in our calculations. As a result, the basic form of source is considered a dipole, and the corresponding magnetic field is the starting point for forward modeling.

Given a magnetic dipole with moment $\vec{m}$ located at the source point $\vec{r}\,'$, the magnetic flux density is given by

$$\vec{B}\left(\vec{r}\right) = \frac{\mu_0}{4\pi}\vec{m}\cdot\nabla\nabla^T\frac{1}{\left|\vec{r}-\vec{r}\,'\right|} = \frac{\mu_0}{4\pi}\mathbf{\Gamma}\,\vec{m} \qquad [74]$$

where μ_0 is the free space permeability and $\mathbf{\Gamma}$ is a dyadic Green's tensor given by the double gradient of the inverse of the distance

between the source and the observation locations. Equation [74] is a slightly different form from that given in the spherical polar coordinates and provides a better implementation for computational purposes. We also note that $\mathbf{\Gamma}$ is also the tensor that defines the gravity gradient tensor produced by a point mass (see Chapter 11.12).

Given a distribution of magnetization as a function of 3-D position, a common approach is to divide the source region into a set of right rectangular prisms whose facets are parallel to the coordinate planes in a Cartesian system. When assuming a constant magnetization within each prismatic cell, then the contribution of one prism to the magnetic anomaly is given by the integration of eqn [74] in the volume of a prismatic cell:

$$\Delta\vec{B}\left(\vec{r}\right)=\frac{\mu_0}{4\pi}\vec{M}\cdot\iiint_{\Delta V}\nabla\nabla^T\frac{1}{\left|\vec{r}-\vec{r}\,'\right|}\mathrm{d}v' \qquad [75]$$

where $\vec{M}$ is magnetization and ΔV is the volume of the prism. The expressions of the integral in eqn [75] can be found in Sharma (1966) and Rao and Babu (1991) for magnetic modeling and in Zhang et al. (2000) for gravity gradiometry modeling.

The equations mentioned previously and other approaches in the literature assume that the magnetization is known and constitute the second step of forward modeling. The magnetization must be defined first before these approaches can be used. As discussed in Section 11.11.2, the magnetization consists of two components, namely, induced and remanent magnetizations, and their vector sum defines the total magnetization.

There are a small number of systems available that can measure three components of the magnetic field vector (Dransfield et al, 2003). Equation [75] yields the required anomaly directly. The majority of modern magnetic data acquired in exploration and environmental applications are the total-field anomaly, which is the difference between the total intensity and the intensity of the Earth's background field:

$$\Delta T=\left|\vec{B}_0+\vec{B}_a\right|-B_0 \qquad [76]$$

where $\vec{B}_0$ and $\vec{B}_a$ are the inducing field and anomalous field, respectively. Under the condition that $B_a \ll B_0$, ΔT is well approximated by the projection of the anomalous field in the direction of the inducing field:

$$\Delta T\cong\vec{B}_a\cdot\hat{B}_0 \qquad [77]$$

Having calculated the anomalous field vector based on eqn [75], the total-field anomaly is readily obtained.

An equally important data type in magnetics is magnetic gradiometry data (Hardwick, 1999; Nelson, 1994; Reford, 2006). Although they are used to a much lesser extent, a significant amount of magnetic gradient data has also been acquired. The acquisition of such data is expected to increase with newer systems available. Two types of magnetic gradient systems are noteworthy. The first is the gradient system based on simultaneous measurements of total-field data using multiple sensors placed at some finite distances. The corresponding data are the finite difference approximation of the gradient of total-field anomaly in eqn [77]. The forward modeling can be carried out analogously by first calculating the total-field anomalies at two sensor locations and then dividing the difference by the sensor separation. A similar type of instrument relies on three-component fluxgate sensors, and the finite difference approach can be used to calculate the magnetic gradient tensor from eqn [76]. The second type of magnetic gradiometry system employs SQUID sensors and measures magnetic gradient tensor directly. Differentiation of eqn [75] is then needed for forward modeling.

11.11.5.2 Forward Modeling of Distributed Sources

Modeling and inversion of exploration-related magnetic data typically involve distributed magnetic sources. One convenient approach is to discretize the source region into a set of continuous right rectangular prisms as discussed in the preceding section. Such an approach is convenient and easy to implement, especially when one deals with only weak magnetic sources without noticeable self-demagnetization effects. For this reason, this approach has been used widely, especially in many inversion algorithms.

This method is conceptually simple and easy to realize but can face huge computational challenges in large-scale problems. Various methods have been developed to accelerate the calculation. These include the use of a wave number-domain approach based on the Fourier expression of the response due to rectangular prisms, wavelet acceleration (see Section 11.11.7), and differential equation approaches.

Here, we briefly describe the wave number-domain approach (e.g., Blakely, 1996; Pilkington, 1997). Let us first examine a single prism centered directly below the origin of a Cartesian coordinate system, with a width of $2a$ and $2b$ in x-direction (northing) and y-direction (easting), respectively, and a depth ranging from h_1 to h_2 (see Figure 22). The Fourier expression of the total-field magnetic anomaly on an observational plane located at vertical coordinate z is given by

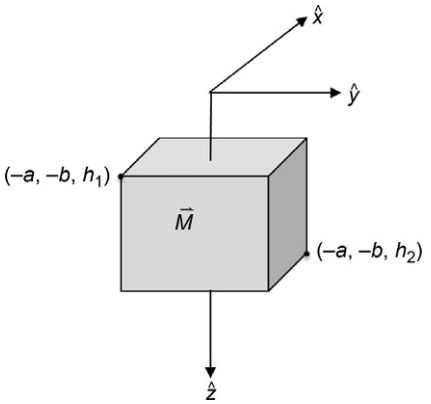


Figure 22 Illustration of the parameters associated with a right rectangular prism located symmetrically below the origin of a Cartesian coordinate system. The prism has a constant magnetization of $\vec{M}$, width of $2a$ and $2b$ in x- and y-directions, respectively, and top and bottom depths of h_1 and h_2. We use a right-hand coordinate system in potential-field methods, where positive x points north, y points east, and z points vertically down.

$$\Delta\widetilde{T}(u,v,z)=\frac{\mu_0 M}{\pi}\left(\hat{M}\cdot\vec{K}\right)\left(\hat{B}_0\cdot\vec{K}\right)\frac{\sin au}{u}\frac{\sin bv}{v}\frac{e^{-(h_1-z)\sqrt{u^2+v^2}}-e^{-(h_2-z)\sqrt{u^2+v^2}}}{u^2+v^2} \quad [78]$$

where u and v are the wave numbers in the x- and y-directions, respectively, $\vec{K}=\left(iu, iv, \sqrt{u^2+v^2}\right)$, $\hat{M}$ is the unit vector of magnetization direction, and $\hat{B}_0$ is the unit vector of the background magnetic field. This is a simple and elegant form in which the direction factor, size factor, and depth factor are clearly outlined. It is interesting to note that the factors associated with the two horizontal dimensions are sinc functions. This is to be expected since the source distribution in the two directions is essentially boxcar functions and the Fourier transform of a boxcar function is a sinc function. The exponentials associated with the vertical dimensions are of course related to the wave number-domain upward continuation operator.

Evaluating this expression, carrying out an inverse Fourier transform yields the desired total-field anomaly in the space domain. The inverse Fourier transform is commonly approximated by 2-D FFT, which works with discrete series and does not involve the actual physical dimensions of the problem. It is necessary to take care of the scaling factor that allows us to approximate the Fourier integral with the discrete Fourier transform. This factor is given by $(2u_N v_N)/\pi$, where u_N and v_N are Nyquist wave numbers in the northing and easting directions, respectively.

Let us now assume that the subsurface source region is discretized into a set of right rectangular prisms using a finite volume-like mesh with a uniform interval in each direction. We further assume each has a known constant magnetization and, without loss of generality, that all cells have the same magnetization direction. Because of the translation-invariant property of the magnetic field, the total-field anomaly due to each horizontal layer is the 2-D convolution of the response of a single prism with the distribution of the magnetization magnitude. Following the convolution theorem, the Fourier transform of the magnetic anomaly due to single layer would be given by the product of the Fourier transform of the magnetization in the layer and the expression in eqn [78]. Summing over the responses of all layers yields the desired magnetic anomaly due to the entire region of the magnetic source:

$$\Delta\widetilde{T}(u,v)=\sum_{k=1}^{M_z}\widetilde{m}(u,v;h_k,h_{k+1})\widetilde{G}(u,v;h_k,h_{k+1}) \quad [79]$$

where $\widetilde{m}(u,v;h_k,h_{k+1})$ is the Fourier transform of the magnetization magnitude by assuming it is located at the center of the prisms and $\widetilde{G}(u,v;h_k,h_{k+1})$ is the expression in eqn [78] with the appropriate physical dimensions of the prism in the layer. Similarly, evaluating eqn [79] and inverse Fourier transforming to the space domain yields the desired modeling result.

In the absence of surface topography and a variable observation height, this is one of the most efficient approaches to performing forward modeling. The efficiency deteriorates when severe surface topography and large relief in observation height are present, because additional space-domain modifications are required to utilize this approach.

11.11.5.3 Full Solution of Induced Magnetization Based on Integral Equations

For a majority of environmental and exploration problems, it is reasonable to assume a linear and isotropic material so that the magnetic susceptibility is a scalar. Magnetization is then given by the product of magnetic susceptibility κ and total magnetic field $\vec{H}$ (see Section 11.11.3):

$$\vec{M}=\kappa\vec{H} \quad [80]$$

where the magnetic field $\vec{H}$ is the total magnetic field and includes both the inducing field $\vec{H}_0$ and the secondary or anomalous field $\vec{H}_a$ produced by the magnetization in the source bodies. The relation between magnetization and susceptibility is therefore nonlinear since the susceptible object has an anomalous field dependent on the magnetization that, in turn, is a function of $\vec{H}$.

A commonly used expression for magnetization in exploration literature is defined directly as $\vec{M}=\kappa\vec{H}_0$ (Blakely, 1996). This is valid when the anomalous magnetic field is much smaller than the inducing field. This condition holds true when $\kappa\ll 1$. In practice, the assumption is valid to a high agree of accuracy when $\kappa<0.1$. For the purpose of forward modeling and inversion, this direct product supplies the magnetization used in eqn [75].

In some geophysical applications, such as mineral exploration for magnetite deposits or environmental cleanup, however, magnetic susceptibility can be much greater, and, as a result, $\left|\vec{H}\right|<\left|\vec{H}_0\right|$, because the secondary field, internal to the magnetic material, opposes the inducing field in the case of paramagnetism. Consequently, the actual magnetization is smaller in magnitude than the approximation mentioned earlier. This phenomenon is termed self-demagnetization. The self-demagnetization effect refers to the extent to which the induced field strength is reduced, due to shape and orientation of the susceptible body relative to the inducing field, by superposition of the secondary field produced by the magnetization inside the body. However, the more important part of the self-demagnetization is the change of magnetization direction away from the inducing field. Because magnetic direction becomes unknown, this aspect of the self-demagnetization effect has a much more negative impact in the interpretation of magnetic data.

Analytic solutions of self-demagnetization effect exist only for limited geometries such as spherical and spheroidal bodies. These solutions have been used to approximate those for irregular-shaped objects. However, because the magnetic field internal to a magnetic body is strongly dependent on its shape, such an approximation leads to substantial errors (Aharoni, 1996; Sharma, 1966). For a finite, nonspheroidal body with high susceptibility, the demagnetization field within the body can also be strongly inhomogeneous. This magnetic phenomenon is more pronounced for bodies with large aspect ratios or highly irregular shapes.

The effect of self-demagnetization has been examined extensively in applied geophysics, especially in mineral exploration (Eskola et al, 1977; Sharma, 1968). One approach to the solution of this problem is to compute demagnetization factors for bodies of simple shapes with analytic solutions such as

those for cylinders (Lee, 1980), spheroids (Traynin and Hansen, 1993), or rectangular prisms (Sharma, 1966). More complicated objects require numerical modeling. Sharma (1966) developed an integral equation approach for this purpose. More recently, Lelievre and Oldenburg (2006) formulated the problem using a differential equation approach. We describe the integral equation approach following Sharma for its simplicity and ease of implementation in practice.

The linear relationship between the magnetization and the magnetic field (eqn [80]) holds true for an isotropic high-susceptibility object situated in the Earth's inducing field, which is a relatively weak field. The total magnetic field within an object is given by

$$\vec{H} = \vec{H}_0 + \vec{H}_a \qquad [81]$$

where $\vec{H}_0$ is the inducing field defined by the IGRF and $\vec{H}_a$ is the secondary field produced by the magnetization $\vec{M}$. The surface measurement of this field constitutes the anomaly. The secondary field is given by

$$\vec{H}_a\left(\vec{r}\right) = \frac{1}{4\pi}\iiint_V \vec{M}\left(\vec{r'}\right)\cdot\nabla\nabla^T\left(\frac{1}{\left|\vec{r}-\vec{r'}\right|}\right)\mathrm{d}v' \qquad [82]$$

where $\vec{r}$ is the position vector of an observation point and $\vec{r'}$ is the position of an arbitrary point inside the source.

Combining eqns [80] and [81] yields the integral equation for the magnetic field:

$$\vec{H}\left(\vec{r}\right) = \vec{H}_0\left(\vec{r}\right) + \frac{1}{4\pi}\iiint_V \kappa\left(\vec{r'}\right)\vec{H}\left(\vec{r'}\right)\cdot\nabla\nabla^T\left(\frac{1}{\left|\vec{r}-\vec{r'}\right|}\right)\mathrm{d}v' \qquad [83]$$

This equation is valid for any observation point $\vec{r}$ in the entire space, including those inside the source body, whereas the integral part concerns only the source region. Knowing the field in the source region allows the calculation of the magnetic field in the whole space. Restricting $\vec{r}$ to be within the volume V, that is, considering $\vec{H}$ on both sides of eqn [83] to be the same function defined inside the source body, yields a Fredholm integral equation of the second kind (Tricomi, 1985).

We note that when the susceptibility is small, the secondary field can be neglected and the total magnetic field $\vec{H}$ in eqn [83] can be approximated by $\vec{H}_0$. This is the standard Born (1999) approximation. Thus, the commonly used expression for magnetization given by the product of the magnetic susceptibility and the Earth's inducing field is an approximation, albeit it is valid to a high degree when susceptibility is <0.1. The so-called self-demagnetization (eqn [83]) is actually the full solution of the magnetic problem.

This integral equation can be solved to obtain the magnetic field $\vec{H}$ inside the source body, which then defines the magnetization as a function of position within the object. For the purpose of modeling magnetic anomaly on the surface, we can proceed with eqn [75].

To solve eqn [83], we can use the method of moments (Harrington, 1968) and a point-collocation scheme. In this approach, we discretize the source body into a set of right rectangular prisms and assume that both magnetic field and magnetization within each cell is uniform and equal to that at the center of the cell. With such a discretization, the continuous integral equation takes on the following matrix form:

$$\vec{H}^m = \vec{H}_0 + \frac{1}{4\pi}\sum_{s=1}^{N}\kappa^s\vec{H}^s\cdot\iiint_{V_s}\nabla\nabla^T\frac{1}{\left|\vec{r}_m-\vec{r}_s\right|}\mathrm{d}v'_s,$$
$$(m,s=1,2,3,\ldots,N) \qquad [84]$$

where $\vec{r}_{ms}$ represents the distance between cells 'm' and 's'; κ^s and $\vec{H}s$ are, respectively, the susceptibility and magnetic field in the cell 's'; and N is the number of cells in the discretized object. Equation [84] represents a linear system of 3 N equations with 3 N unknowns since there are three magnetic field components for each cell. The resulting matrix system can be solved using a suitable linear solver (Barrett et al., 1993). The solution yields the values of the magnetic field components, and the magnetization, of all the cells in the source body.

11.11.5.4 Solution of Induced Magnetization Based on Differential Equations

Historically, the magnetic problem is mostly considered in the framework of integral formulation as shown in eqns [75] and [83] in applied geophysics. However, the problem is governed by differential equations, namely, Maxwell's equations. Therefore, it is logical to solve the magnetic forward modeling problem using the differential equations. For large-scale problems, differential equation solutions offer certain advantages because of the inherently sparse numerical systems involved and the availability of efficient iterative solvers developed in numerical analysis community.

Lelièvre (2005) and Lelievre and Oldenburg (2006) applied this approach to the solution of self-demagnetization problem. Davis et al. (2013) applied this approach in the inversion of regional-scale magnetic data. It is expected that this route of magnetic modeling will gain much more utility in geophysics. We provide a brief description of the method here. Readers are referred to the references mentioned earlier for more details.

Consider the Maxwell's equations. The magnetic flux density $\vec{B}$ is divergence-free:

$$\nabla\cdot\vec{B} = 0 \qquad [85]$$

In the absence of free current as the magnetic source, the static magnetic field $\vec{H}$ is curl-free:

$$\nabla\times\vec{H} = \vec{0} \qquad [86]$$

The two fields are related by the constitutive relationship:

$$\vec{B} = \mu\vec{H} = \mu_0(1+\kappa)\vec{H} \qquad [87]$$

where μ_0 is the free space permeability and κ is the magnetic susceptibility, which is a function of spatial position. Because of eqn [86], we can define a scalar potential ϕ_H such that the magnetic field is given by its negative gradient:

$$\vec{H} = -\nabla\phi_H \qquad [88]$$

Substituting into eqn [85] yields the well-known static elliptical equation:

$$\nabla\cdot(\mu\nabla\phi_H) = 0 \qquad [89]$$

This equation is well studied in geophysics. For instance, the DC resistivity problem is governed by this equation, except for the nonzero source term on the right-hand side (Dey and Morrison, 1979). Correspondingly, many efficient numerical solution strategies are available. Dey and Morrison (1979) and Lelièvre (2005) both used the finite volume method. The resulting linear system of equation is sparse, symmetrical, and positive definite. A preconditioned conjugate gradient (CG) method can be used to obtain the solution efficiently.

The remaining issues are the boundary conditions. The finite volume mesh must be extended to large distances in all directions away from the magnetic regions. The mesh also includes the region of air above the Earth's surface. When the boundary distance is sufficiently large, the anomalous field produced by the magnetization associated with the susceptibility distribution κ is negligibly small. The gradient of the magnetic potential to be solved for should then be equal to the Earth's background field. This is one of the direct means to specify the boundary conditions. Other sophisticated methods are available to better approximate the boundary conditions while allowing a less stringent requirement of the distance between the mesh boundary and the anomalous susceptibility region.

11.11.5.5 Numerical Example

As an illustration, we model the magnetic response of the two highly susceptible metallic objects from UXO cleanup (Sanchez et al., 2008). The first object is an 81 mm M57 projectile and the second object is an irregularly shaped fragment (**Figure 23**). The projectile's longest axis and fragment's plane both lie on the horizontal plane. The projectile's tapered endpoints towards south and the fragment's two arms point north and east, respectively. We have modeled the magnetization distribution in each object using the integral equation solution in 84. The inducing field magnitude is 50 000 nT with an inclination of 65° and declination of 25°. We note that the magnetization is approximately parallel to the axis of the 81 mm projectile. In the case of the fragment, the magnetization is rotated from the inducing field to the long axis of each arm.

Figures 24 and **25** compare responses of the magnetic anomalies of the projectile and fragments, respectively. The observation plane is 0.80 m above the top of the models. Plots of tensor components are arranged according to the order in the gradient tensor. The total-field anomaly is plotted in the lower left corner of the figures. The gradient maps in each figure have individual color bars. The anomalies are calculated by summing the fields from all cells in each model. Both data sets are strongly dipolar. It is important to note that, if the magnetic body is a sphere, the dipole axis will be oriented in the direction of the inducing field. For other object shapes, the dipole axis tends to rotate towards the object's longest axis. By following the line from peak to trough in these plots, it is possible to infer a dipole oriented approximately in the north–south direction for the 81 mm projectile and a dipole oriented in the northeast direction for the fragment. Of course, detailed examination of the plots also indicates the presence of nondipolar components.

11.11.6 Equivalent Source Techniques

Fourier transform-based processing techniques have played an important role throughout the history of potential-field methods. The use of FFT-based processing has become ubiquitous and almost synonymous to potential-field processing. However, the applicability of FFT processing is predicated on two major assumptions. The first and perhaps the most important is that the observations are on planar surfaces, and the second is that the data have been interpolated to a uniform grid using an adequate gridding algorithm that preserves the harmonic property of the data. However, the majority of data in exploration or crustal studies do not satisfy these assumptions, especially the planar surface assumption. The violation of these assumptions renders FFT-based processing invalid. The equivalent source technique first proposed by Dampney (1969) offers a much needed and powerful alternative to the FFT approach.

The equivalent source technique is more general and does not rely on the two assumptions mentioned earlier. Consequently, it plays an important role in the processing and interpretation of potential-field data in general and in magnetic data processing in particular. The important areas of application are interpolation of scattered data (Cordell, 1992; Mendonça and Silva, 1994), processing of data acquired on uneven surfaces (Hansen and Miyazaki, 1984; Nakatsuka and Okuma, 2006), RTP of magnetic data at low latitudes (Li and Oldenburg, 2000a,b; Silva, 1986), conversion between different components of potential-field data (Emilia, 1973; Li and Li, 2014), denoising potential-field data (Li, 2001; Salem et al., 2010), and joint processing of multiple data sets (Davis and Li, 2012). This technique involves constructing a fictitious source layer located below the observation surface that reproduces the observed data and then computing, from the constructed source layer, the desired field above the equivalent source.

The equivalent source technique stems from the harmonic properties of the potential fields in the source-free region. If the potential is known on the bounding surface of the region, then the field anywhere in this region is uniquely determined by the solution of a Dirichlet boundary value problem (Kellog, 1953). Therefore, if one can construct a source distribution outside this region that reproduces the field on the boundary, then it also defines the field in the source-free region. There are infinitely many such source distributions, but only one of them need be constructed. Since it usually bears no resemblance to the true source, the term equivalent source is used to describe it.

In the original method proposed by Dampney (1969) and in many follow-up works, one would choose a set of locations for the placement of discrete equivalent sources, often tied to the observation locations. Different types of equivalent sources have been employed such as monopoles and point dipoles. Correspondingly, an even-determined linear system of equations is formed. The inversion of such a system yields the desired equivalent source distributions. More recent work has extended the method by utilizing continuous source distributions to avoid the difficulties associated with the instability of discrete point sources (Hansen and Miyazaki, 1984). Furthermore, the influence of applied geophysical inversion has steered the method away from the exact fitting of observation and towards fitting the data according to the statistics of the

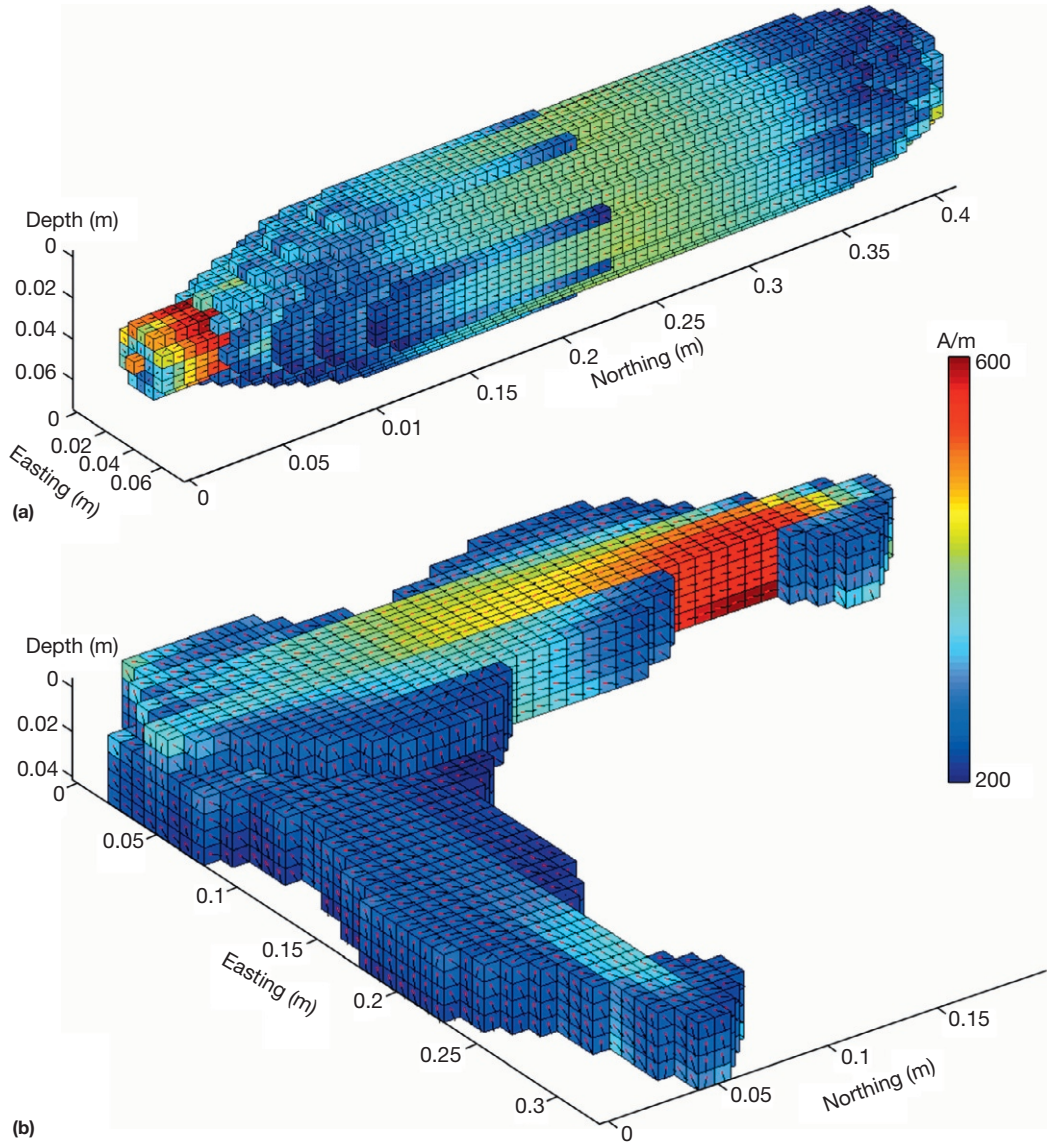


Figure 23 Examples of unexploded ordnance (UXO) modeling. The top object is an 81 mm projectile and the bottom one is a simulated fragment. The color shading and arrows on the cell facets show the strength and directions of magnetization obtained by solving the full self-demagnetization using eqn [84]. The inducing field is in the direction of (65°, 25°), yet the magnetization is approximately parallel to the axis of the 81 mm projectile. In the case of the fragment, the magnetization is rotated from the inducing field to the long axis of each arm.

data error (Davis and Li, 2011; Li and Oldenburg, 2010; Oliveira et al., 2013). We describe the inverse formulation of equivalent source technique in the succeeding text.

11.11.6.1 Inverse Formulation of Equivalent Source Construction

Let us assume a continuous distribution of equivalent source functions on the chosen surface located below the observational surface, instead of using point sources. The source is then related to the field by a Fredholm integral equation of the first kind. For magnetic data processing, a natural and effective choice is the magnetization aligned in the inducing field direction.

Let $\vec{d}$ be the potential-field data located over an observation surface, S_o, and let $m\left(\vec{r}\right)$ be the strength of the equivalent source distributed over a surface S_e (**Figure 26**). The surface S_e lies below S_o or coincides with it but must be above actual sources that produced the data, $\vec{d}$. The integral equation is formally written as

$$d\left(\vec{r}\right) = \iint_{S_e} m\left(\vec{r}\,'\right) g\left(\vec{r}\,', \vec{r}\right) \mathrm{d}s', \quad \vec{r} \in S_o \qquad [90]$$

where $g\left(\vec{r}\,', \vec{r}\right)$ is the kernel function, which here is given by the function that relates the magnetization with a known direction to the total-field anomaly.

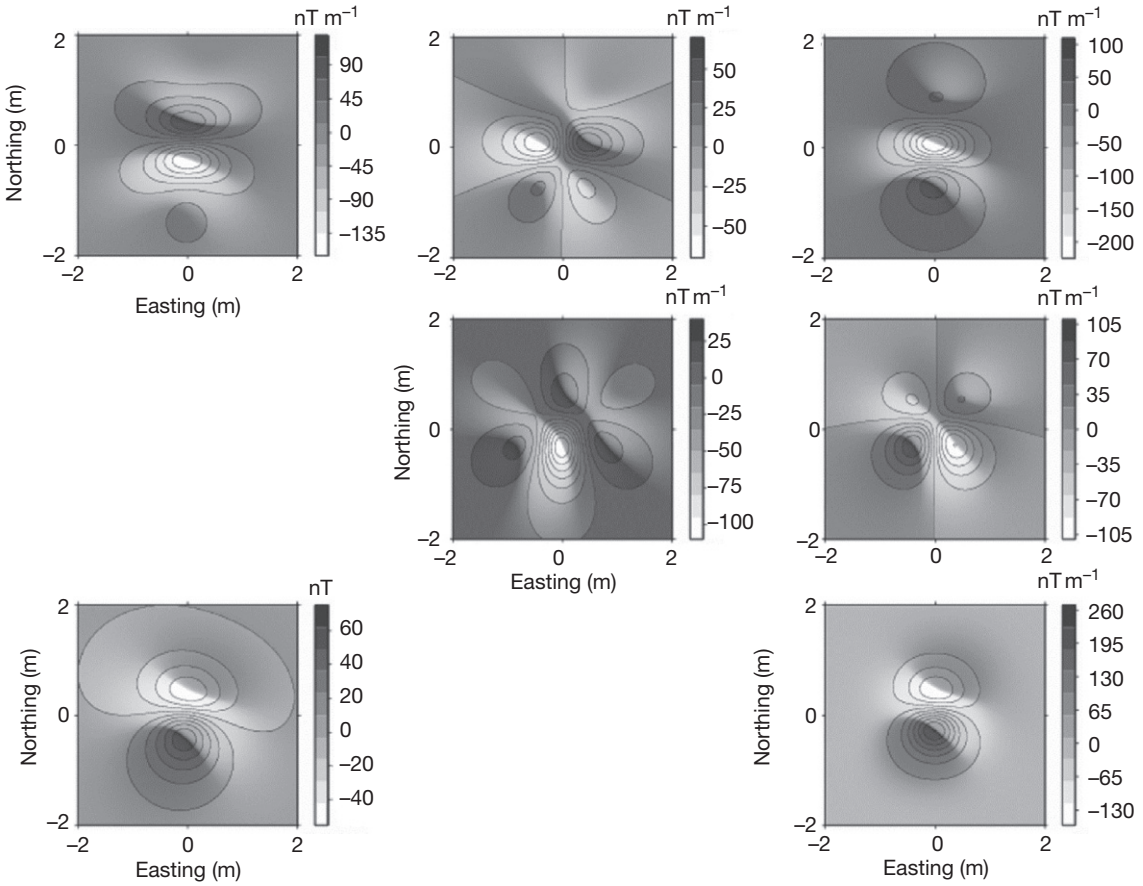


Figure 24 Total-field anomaly and magnetic gradient tensor produced by the 81 mm projectile in **Figure 23(a)**. The total-field anomaly is displayed in the lower left panel. The gradient tensor components are shown according to their position in the tensor. For instance, top left panel is the T_{xx} component and lower right panel is the T_{zz} component. Reproduced from Sanchez V, Li Y, Nabighian M, and Wright D (2008) Numerical modeling to identify relative contributions of high-order magnetic moments in UXO discrimination. *IEEE Transactions on Geoscience and Remote Sensing* 46: 2568–2583.

The construction of the equivalent source requires the inversion of this equation to recover $m\left(\vec{r}\right)$ from the given data, $\vec{d}$. For a numerical solution, this entails the discretization of eqn [90]. We divide the undulating surface S_e into a set of contiguous rectangular cells and assume a constant value of magnetization strength within each cell. All cells are assumed to have the same magnetization direction. Such a discretization then assumes a piecewise constant equivalent source and yields a matrix representation:

$$\mathbf{G}\,\vec{m}=\vec{d} \qquad [91]$$

where $\vec{d}=(d_1,\cdots,d_N)^T$ is the vector containing the observed data, $\vec{m}=(m_1,\ldots,m_M)^T$ is the vector containing the strength of the equivalent source, and G is the coefficient matrix whose elements g_{ij} are typically defined by

$$g_{ij}(\vec{r}_i)=\iint_{\Delta S_{ej}} g\left(\vec{r}\,',\vec{r}_i\right)\mathrm{d}s' \qquad [92]$$

Equations [91] and [92] define the forward modeling in classical linear inverse problem. The construction of the source distribution then follows the Tikhonov inversion in a straightforward manner.

Let φ_d be the L_2 measure of the data misfit function and φ_m be the model objective function:

$$\varphi_m=\iint\left\{\alpha_s m^2+\left(\frac{\partial m}{\partial x}\right)^2+\left(\frac{\partial m}{\partial y}\right)^2\right\}\mathrm{d}x\mathrm{d}y \qquad [93]$$

which can be represented in a matrix format $\vec{m}\,TW_m^TW_m\,\vec{m}$ using the same discretization for generating eqn [91]. The corresponding solution is obtained by minimizing a weighted sum of the data misfit and model objective functions by solving the following system of equations:

$$\left(G^TG+\beta W_m^TW_m\right)\vec{m}=G^T\,\vec{d} \qquad [94]$$

where β is the regularization parameter to be found during the construction process so that the data misfit function attains a value consistent with the error level in the data.

Once the equivalent source is obtained, it can be used to calculate the field itself or any related components in the 3-D space above the observation surface. We will illustrate using examples later.

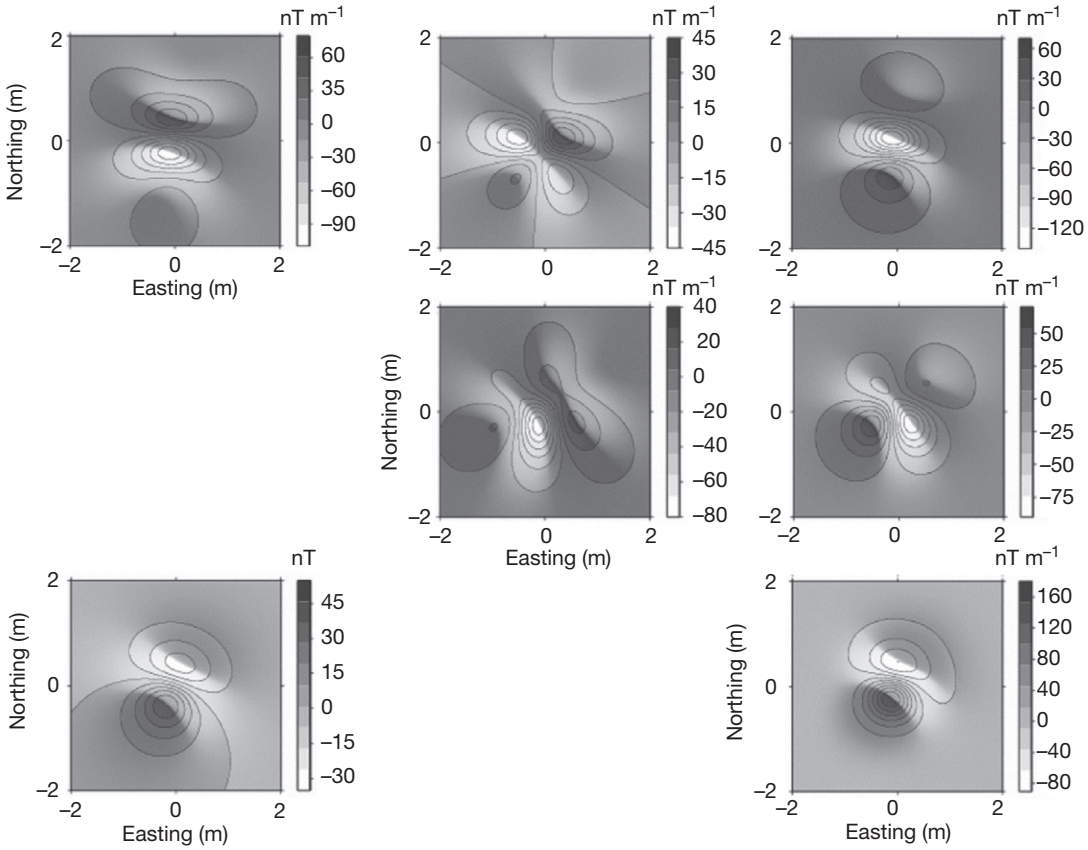


Figure 25 Total-field anomaly and magnetic gradient tensor produced by the fragment in **Figure 23(b)**. The total-field anomaly is displayed in the lower left panel. The gradient tensor components are shown according to their position in the tensor. Reproduced from Sanchez V, Li Y, Nabighian M, and Wright D (2008) Numerical modeling to identify relative contributions of high-order magnetic moments in UXO discrimination. *IEEE Transactions on Geoscience and Remote Sensing* 46: 2568–2583.

Several issues arise in practice. Algorithmically, the regularization parameter is selected such that the data misfit value for the constructed equivalent source is equal to the expected value of the misfit function. This assumes that the error statistics are known. However, the noise statistics are often unknown. Therefore, different criteria commonly used in linear inverse theory must be employed to estimate the regularization parameter. This process also yields an estimate of the noise in the data, and the corresponding data reproduced by the equivalent source are considered to be a denoised version of the observations (Li, 2001). Two commonly used criteria are the *L*-curve (Hansen, 1992) and generalized cross validation (Wahba, 1990).

Computationally, the solution of eqn [94] can be prohibitively expensive. This challenge has been discussed by many authors. However, increased computer power and algorithmic improvements are beginning to overcome these difficulties, and applicability of the equivalent source method is improving. The challenge lies in the fact that the coefficient matrix is dense. For large problems, the need to store and apply the coefficient matrices is a major impediment to solving large-sized equivalent construction problems in a realistic time frame. Several approaches have been developed in the literature to speed up the process. The size of the dense matrix G is determined by the number of data N and the number of cells M in the discretized equivalent source layer. Intuitively, any effort to reduce these would be fruitful. For example, Mendonça and Silva (1994) developed a procedure to adaptively add more data to the construction in order to limit the computational cost. Ridsdill-Smith and Dentith (1999) used wavelets along profiles of magnetic data in order to find edges in the equivalent source so that no unnecessary parameters are used. Li and Oldenburg (2010) assumed gridded data and solved the

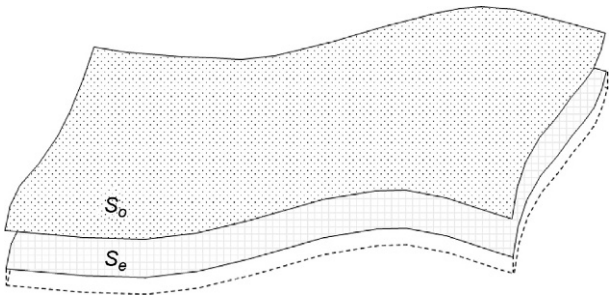


Figure 26 Conceptual setup of equivalent source construction. S_o is the observation surface and S_e represents the top surface of the equivalent source.

construction problem in the wavelet domain by forming a sparse representation of the matrix in the same domain.

For general applicability, one can combine adaptive discretization of the equivalent source with the matrix compression to reduce the storage requirement and the computational time during the construction (Davis and Li, 2011). In this approach, one uses a quadtree mesh discretization to form a semistructured piecewise representation of the source so that only fine discretization is used where rapid changes occur in the data and course discretization is used where the observed data vary slowly (see **Figure 27**). The use of an adaptive quadtree mesh can yield a factor of 4–6 in the reduction of the number of cells to represent the equivalent source.

The cells in the semistructured mesh are then ordered using the Hilbert space filling curve (Hilbert, 1891; Moon et al., 2001) to form an algebraic vector in such a way that any two elements that are adjacent to each other in the algebraic vector are also close to each other in the space domain of the equivalent source (see **Figure 28**). This ordering then allows one to compress efficiently each row of the coefficient matrix G using a 1-D wavelet transform based on compactly supported orthonormal wavelets (e.g., Daubechies, 1992; Mallat, 1989) and subsequent thresholding of the wavelet coefficients. The wavelet compression works by discarding zero or near-zero coefficients and storing only the significant wavelet transform coefficients, which reduces the required memory for storing the matrix. The matrix multiplications $\left(G^T G\,\vec{m}\right)$ are then carried out in the wavelet domain by working with only these large coefficients. The ability to work with a smaller number of coefficients achieves the needed computational efficiency. This step can achieve a reduction in storage and required computation time by a factor of more than 100.

In summary, the approach consists of three components of adaptive quadtree discretization, ordering by Hilbert space filling curve, and 1-D wavelet transform-based compression. To our knowledge, this is one of the most efficient methods for equivalent source construction. More details can be found in (Davis and Li, 2011) and in a parallel discussion on a similar approach applied to the 3-D inversion of magnetic data.

11.11.6.2 Examples

11.11.6.2.1 A synthetic example

We now illustrate the utility of the equivalent source methods using examples from three different areas of application: conversion to different components and RTP on uneven surfaces. We note that the results are produced using variations of the approaches described earlier. For brevity, we omit the description of the details of the example. Instead, we focus on the fact that different approaches can be used to construct an equivalent source and such a source can be used to process magnetic data in cases where wave number-domain approaches are not applicable.

Figure 29 displays a set of synthetic total-field data simulated on an undulating observation surface. The total-field data have been contaminated by Gaussian noise with 2 nT standard deviation. The observation surface consists of a valley and a hill with a total relief of 150 m. We note that this relief is not

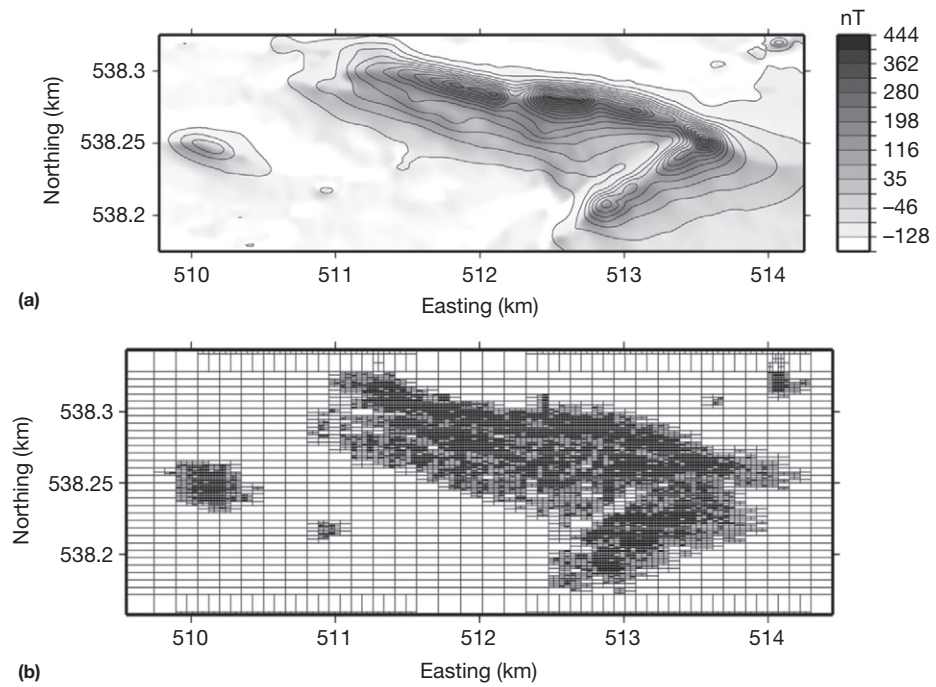


Figure 27 An example of adaptive quadtree mesh for processing magnetic data using an equivalent source. Top panel shows the total-field magnetic anomaly data. Both the magnitude of the anomaly and its spatial rate of change vary significantly within the data set. Lower panel displays corresponding quadtree mesh that defines an equivalent source mesh. Finer discretization is used where high magnetic anomaly and steep gradients are present. Reproduced from Davis K and Li Y (2011) Fast solution of geophysical inversion using adaptive mesh, space-filling curves, and wavelet compression. *Geophysical Journal International* 185: 157–166.

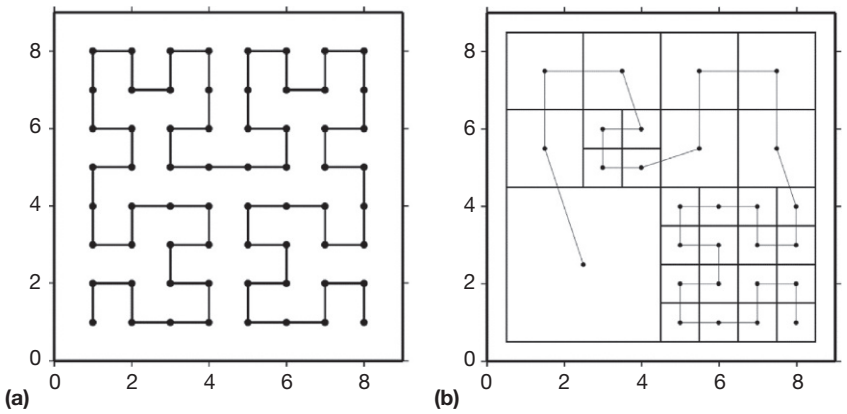


Figure 28 Illustration of ordering by Hilbert space filling curve applied to a regular 2-D grid (a) and to a quadtree adaptive mesh (b). The curve connecting cell centers indicates the ordering starting from lower left corner. The aim of the ordering by Hilbert space filling curve is to represent the cells in the equivalent source in an algebraic vector such that any two elements that are adjacent to each other in the algebraic vector are also close to each other in the equivalent source. Reproduced from Davis K and Li Y (2011) Fast solution of geophysical inversion using adaptive mesh, space-filling curves, and wavelet compression. *Geophysical Journal International* 185: 157–166.

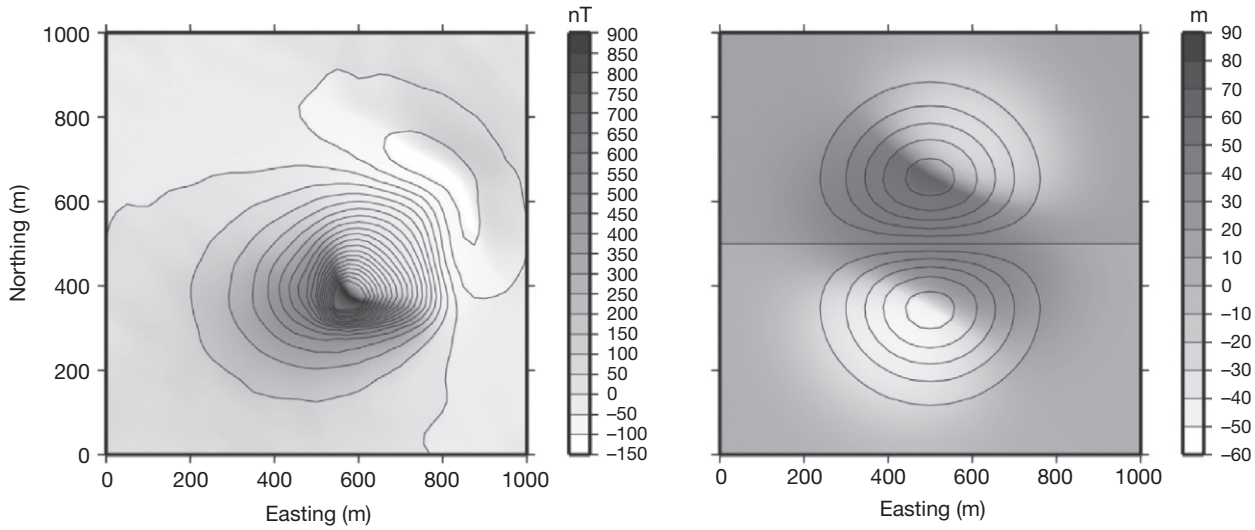


Figure 29 A synthetic example to illustrate the utility of equivalent source technique for processing magnetic data. On the left is the total-field anomaly observed on a topographic surface. The ambient field has an inclination of 65° and a declination of 25°. On the right is the elevation of the simulated observation surface. The total relief is 150 m.

extreme over the distance of 1 km. Even with such a height variation, the planar surface assumption necessary in FFT processing in the wave number domain is violated. We process this data set using both equivalent source and wave number-domain techniques and compare the results.

The construction of the equivalent source fits the data to the level of the known noise standard deviation. The resulting equivalent source is then used to compute desired data such as the three orthogonal components or the RTP field. The equivalent source construction takes into full account the variation in observation height. The wave number-domain approach directly applies the operators without additional processing or filtering and assumes that the data are on a planar surface. No additional processing or filtering is applied so as to ensure a comparison that highlights the essential issues without unnecessary complications.

Figure 30 shows the comparison between the results from the two techniques when performing component conversions. For brevity, we only display the x-component. The top row shows B_{ax} derived from the equivalent source method, and the corresponding difference between the true B_{ax} computed from the original synthetic model and the same quantity converted from the noisy total-field anomaly in Figure 29. The bottom row displays the results from the wave number-domain approach. We first note that there is a large difference between the two versions of computed x-component, which is a direct consequence of neglecting the observation height in the wave number-domain approach. This conclusion is supported by the differences from the true B_{ax} values. The difference for the equivalent source result is mostly confined within 5 nT, with a few isolated points approaching 20 nT. The distribution appears to be random spatially and symmetrical about

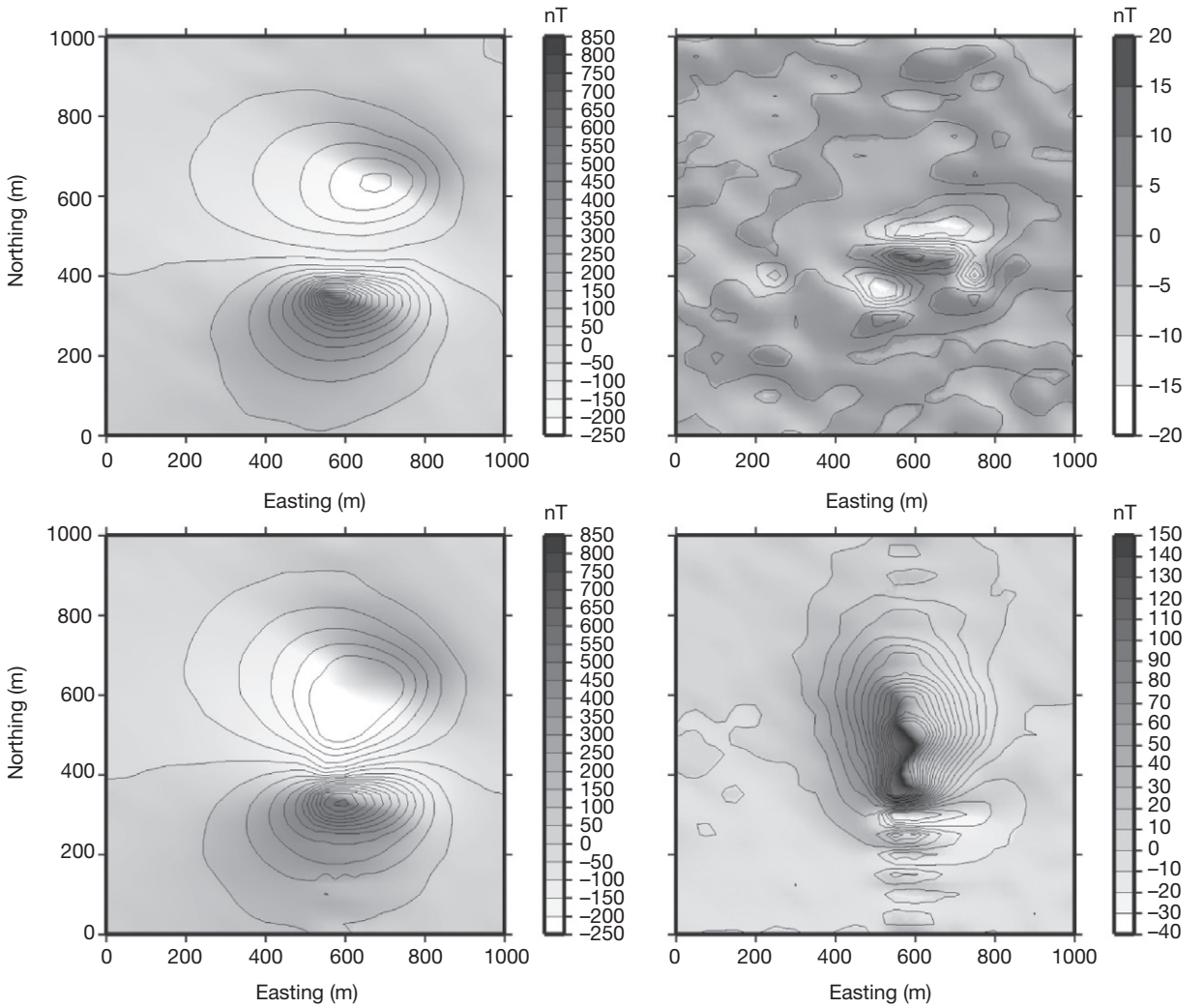


Figure 30 Comparison of component conversion using equivalent source construction and wave number-domain operator. Top row shows the *x*-component of magnetic anomaly (left panel) computed from the equivalent source method and the corresponding difference (right panel) from the true field. The equivalent source construction takes into account the observation height of each datum. Bottom row shows the converted x-component from the wave number-domain operator (left) by assuming a planar observation surface and the difference (right) from the true field.

zero. The difference for the wave number-domain result, however, reaches as high as 150 nT and shows clear coherent spatial structure.

One area directly impacted by this contrast in component conversion is in the amplitude inversion (Section 11.11.8) in the presence of remanent magnetization. The amplitude data must be computed from the observed total-field anomaly by first converting to three orthogonal components. The large error produced by the wave number-domain approach because of large variations in observation height is unacceptable. The equivalent source technique provides a reliable means for that conversion.

Figure 31 shows the comparison between the results from two techniques when performing RTP operation. The results are displayed in the same format as Figure 30: Top row is the result from the equivalent source technique and the bottom row corresponds to the wave number-domain result. The contrast is as pronounced as in the case of component conversion.

11.11.6.2.2 A field example

The flexibility of equivalent source processing also allows one to work with multiple data sets. These data could be of the same type but acquired at different times, or they can be different types of data. For example, different total-field data acquired from different surveys with different observation locations in 3-D can be directly processed together. Total-field anomaly and magnetic gradient can also be processed together. Figure 32 shows the result of jointly processed total-field anomaly and gradient data by combining them as independent but complementary data sets to derive a high signal-to-noise data set. A single equivalent source is constructed to reproduce total-field data and crossline gradient measured using multiple magnetometers in the central Abitibi Destor-Porcupine–Pipestone Fault area, Ontario (Ontario Geological Survey, 2004). The three radial power spectra shown are, respectively, the total-field data gridded using minimum curvature, total-field data processed using equivalent

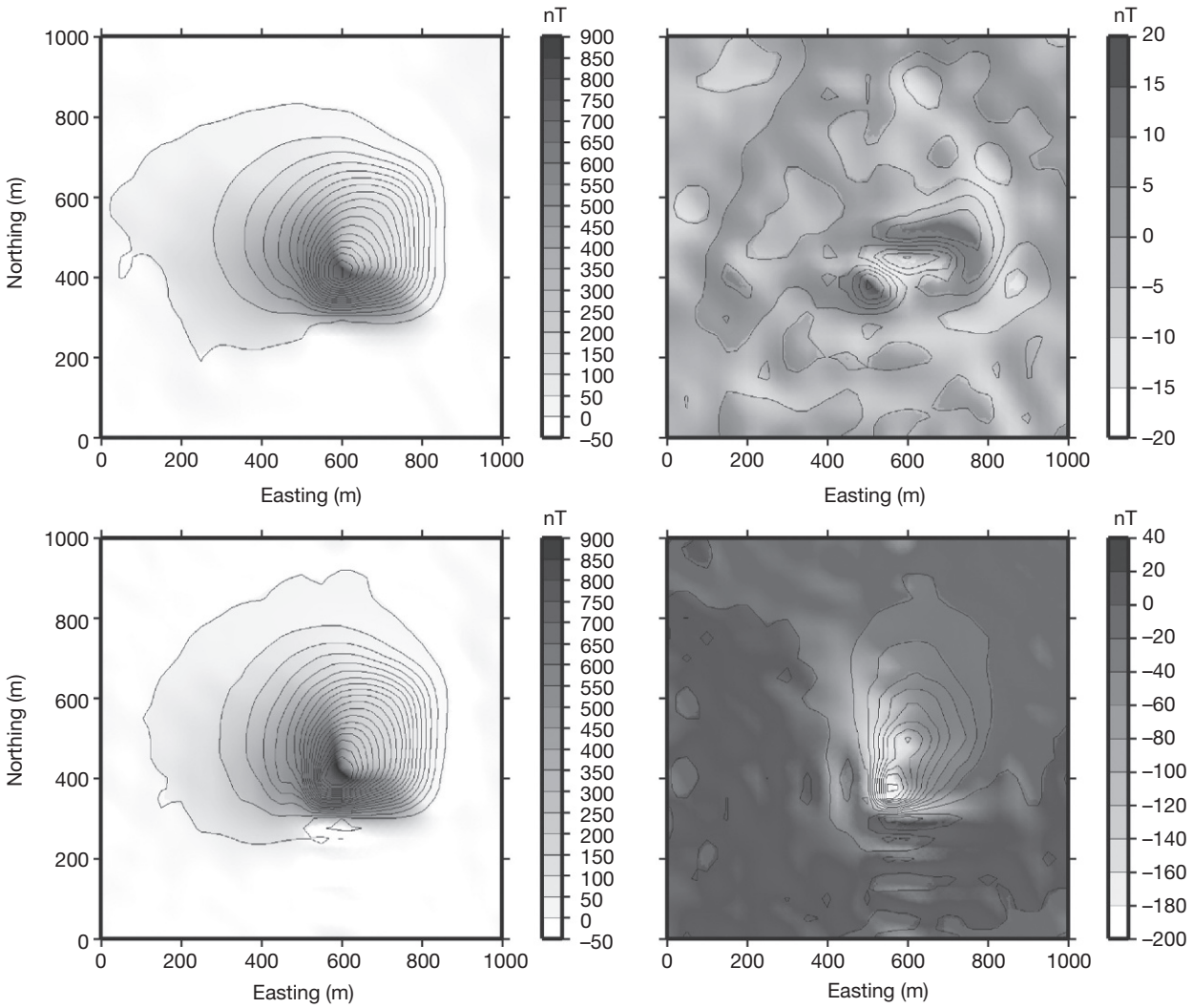


Figure 31 Comparison of reduced-to-the-pole results using equivalent source construction and wave number-domain operator. Top row shows the RTP field of magnetic anomaly computed from the equivalent source method (left) and the corresponding difference from the true field (right). Bottom row shows the RTP field from the wave number-domain operator (left) and the difference from true RTP field (right).

source, and total-field and gradient data jointly processed using an equivalent source (Davis and Li, 2012). The effect of denoising through equivalent source methods is shown by the significant decrease in high frequency content compared with the minimum curvature data. Furthermore, an increase of power in the mid-frequencies (corresponding to wavelengths between 80 and 250 m) is also noticeable for the jointly processed data. This increase essentially states that the extra information brought in by the gradient data is in this wavelength band.

11.11.7 Magnetic Susceptibility Inversion

Inversion is a well-established tool for assisting in the interpretation of geophysical data by constructing a representation of the subsurface physical property distribution or by determining the geometrical boundaries of geologic structures from the observed data. This approach has played an important role in magnetic methods. Similar to the case in other geophysical methods, the use of inversion in magnetics has led to two major paradigm shifts: (1) moving interpretation from the data domain to the model domain where one has a representation of the geology defined in 3-D space and (2) shifting the information extraction from qualitative to quantitative in nature. The use of inverse methods was originally focused on anomaly-scale interpretation, but the availability of generalized inversion has been steadily expanding the use to regional-, continental-, and even global-scale problems.

Inverse formulations in magnetic methods can be classified into three general categories. The first is the parametrical inversion for the geometry of isolated causative bodies and the associated physical properties such as magnetic susceptibility or magnetization. A parametric inversion is carried out to recover the parameters of simple targets through the use of nonlinear least squares (e.g., Ballantyne, 1980; Bhattacharyya, 1980; Silva and Hohmann, 1983; Whitehill, 1973). Alternatively, causative bodies are represented as polygonal bodies in

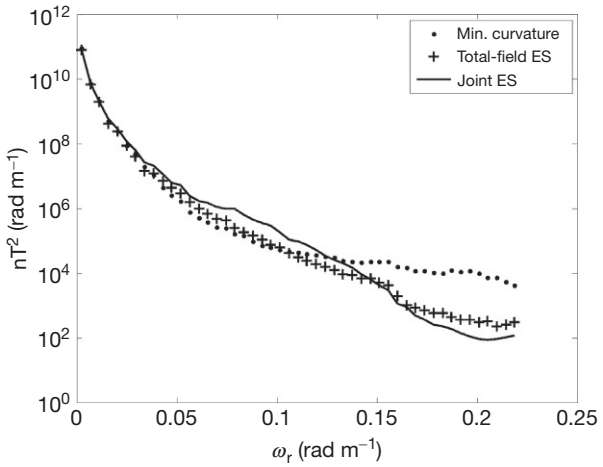


Figure 32 The radial power spectral density plots of the observed total-field anomaly (dots), data created by total-field equivalent sources (pluses), and data modeled by the joint total-field and gradient equivalent sources (solid line). The use of gradients increases power of the signal in mid-frequencies. Additionally, high-frequency noise is suppressed greater than the total-field equivalent source method or minimum curvature gridded data. Reproduced from Davis K and Li Y (2011) Fast solution of geophysical inversion using adaptive mesh, space-filling curves, and wavelet compression. *Geophysical Journal International* 185: 157–166.

2-D or polyhydronal bodies in 3-D (Pedersen, 1979; Wang and Hansen, 1990), and the vertices of the objects are recovered as the unknowns. The second category is represented by the generalized geometrical inversion for basement depth (e.g., Barbosa and Silva, 2011; Martins et al., 2011; Pedersen, 1979; Pilkington and Crossley, 1986). The third category consists of generalized physical property inversions for either the magnetic susceptibility or the magnetization distribution in 3-D space. Parker and Huestis (1974) in their crustal studies inverted for the distribution of magnetization in a layer. Cribb (1976) represented the magnetic source by a set of dipoles and attempted to recover the dipole strength by applying linear inverse theory. Guillen and Menichetti (1984) chose to use a prismatic representation and performed a regularized inversion by minimizing the moment of inertia of the causative body. Li and Oldenburg (1996, 2000, 2003) formulated a generalized 3-D inversion of magnetic data by using the Tikhonov regularization and a model objective function that measures the structural complexity of the model and incorporates a depth-weighting function (see succeeding text). In this section, we focus on the last category because of its widespread use in the exploration and solid Earth studies.

When inverting for magnetic properties, we implicitly assume that the subsurface geology is represented by the variation of the magnetic property being inverted for. The representation is commonly referred to as a model and it attributes magnetic information to each spatial point in the subsurface. Earlier development of magnetic inversion focused on susceptibility, while later development extended to magnetization that consists of both induced and remanent magnetizations. In this section, we refer to the inversions based on these two model representations, respectively, as susceptibility inversion and magnetization inversion. We discuss the fundamentals of susceptibility inversion in this section and apply the formalism to the magnetization inversion in the next section.

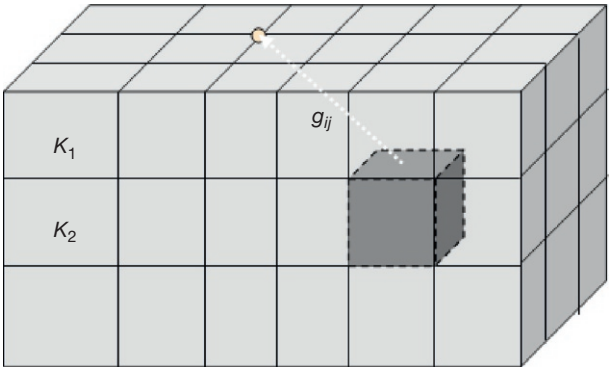


Figure 33 Discretization of subsurface volume into a set of contiguous rectangular cells. Each cell is assumed to have a constant magnetic susceptibility.

11.11.7.1 Susceptibility Inversion

In many cases, one can assume that there is no remanent magnetization and the self-demagnetization effect can be neglected. Consequently, the direction of magnetization is assumed to be the same as the current inducing field direction. This is a valid assumption in a majority of cases, as evidenced by many successful applications. Under such assumptions, the primary physical property in question is the magnetic susceptibility, κ, which determines the magnetization when subjected to the Earth's inducing field. As discussed in the modeling section, the induced magnetization without self-demagnetization effect is defined by (Blakely, 1996)

$$\vec{M} = \kappa \vec{H}_0 = \kappa \vec{B}_0 / \mu_0 \qquad [95]$$

The corresponding magnetic anomaly vector $\vec{B}_a$ a point $\vec{r}$ is given by

$$\vec{B}_a\left(\vec{r}\right) = \frac{1}{4\pi}\iiint_V \kappa\left(\vec{r}\,'\right)\vec{B}_0 \cdot \nabla\nabla\left(\frac{1}{\left|\vec{r} - \vec{r}\,'\right|}\right) \mathrm{d}v' \qquad [96]$$

Projection of this anomalous vector onto a given direction yields a particular component of the anomaly, $\Delta B_u = \vec{B}_a \cdot \hat{u}$. For example, $\hat{u} = \hat{z}$ yields the vertical component and $\hat{u} = \hat{B}_0$ yields the total-field anomaly ΔT when the anomalous field is much smaller than the background field at the observation location:

$$\Delta T\left(\vec{r}\right) = \frac{1}{4\pi}\iiint_V \kappa\left(\vec{r}\,'\right)\vec{B}_0 \cdot \nabla\nabla\left(\frac{1}{\left|\vec{r} - \vec{r}\,'\right|}\right) \cdot \hat{B}_0 \mathrm{d}v' \qquad [97]$$

Thus, the total-field anomaly is linearly related to the subsurface magnetic susceptibility. It follows that classical linear theory can be used to formulate an inversion algorithm to recover the magnetic susceptibility.

To perform an inversion, we must first discretize the subsurface volume of interest. The easiest approach is to divide the volume into a set of contiguous rectangular prisms (Figure 33).

The size of the cells is chosen to be smaller than the resolvable scale for the data set to be inverted, so that we can assume each cell has a constant magnetic susceptibility. Each cell would have a contribution to the datum at each observation location:

$$\Delta T_i = \sum_{j=1}^{M} g_{ij}\kappa_j,\ i = 1, \cdots, N \qquad [98]$$

where N is the number of data, M is the number of cells, and g_{ij} is Green's function representing the contribution of the jth cell at the ith location for unit susceptibility

$$g_{ij} = \iiint_{\Delta V_j} \frac{1}{4\pi} \vec{B}_0 \cdot \nabla\nabla \left(\frac{1}{\left| \vec{r} - \vec{r}\,' \right|} \right) \cdot \hat{B}_0 \mathrm{d}v' \qquad [99]$$

This relationship in eqns [98] and [99] can be expressed in a compact matrix form as

$$\vec{d} = \mathbf{G}\,\vec{\kappa} \qquad [100]$$

where $\mathbf{G}$ is the sensitivity matrix, $\vec{d}$ the data vector and, and $\vec{\kappa}$ the model vector.

Once the problem is discretized and the forward modeling eqn [100] is defined, we can define the major components of the inversion algorithm of which there are three. The first component is the data misfit function:

$$\begin{aligned} \varphi_d &= \sum_{i=1}^{N} \left(\frac{\Delta T_i^{\mathrm{obs}} - \Delta T_i^{\mathrm{pre}}}{\sigma_i} \right)^2 \\ &= \left(\vec{d}^{\,\mathrm{obs}} - \vec{d}^{\,\mathrm{pre}} \right)^T W_d^T W_d \left(\vec{d}^{\,\mathrm{obs}} - \vec{d}^{\,\mathrm{pre}} \right) \end{aligned} \qquad [101]$$

where $\vec{d}^{\,\mathrm{obs}} = \left(\Delta T_1^{\mathrm{obs}}, \ldots, \Delta T_N^{\mathrm{obs}} \right)^T$ are the observed total-field data, $\vec{d}^{\,\mathrm{pre}} = \left(\Delta T_1^{\mathrm{pre}}, \ldots, \Delta T_N^{\mathrm{pre}} \right)^T$ are the predicted data, σ_i is the standard deviation of the error associated with the ith datum, and $W_d = \mathrm{diag}\{\ldots, 1/\sigma_i, \ldots\}$.

The data misfit function is one of the key components because the recovered model must reproduce the observed data and the data misfit function provides the measure. The final data misfit must be sufficiently small. Any model that does not satisfy this criterion could not represent the geology that has produced the observation. However, the data misfit should not be zero either. This is because all data contain errors and a zero misfit means the recovered model is reproducing the noise and the data.

The optimal data misfit is the statistical expectation of the data misfit function. Assuming that the composite errors in the data have a normal distribution, then the misfit function in eqn [101] is a random variable with a χ^2 distribution and its expectation is equal to the number of data, N.

The second component of the inversion is a model objective function that measures the complexity of the recovered susceptibility distribution:

$$\begin{aligned} \varphi_m(\kappa) = {} & \alpha_s \int_V w_s \left[w(z)(\kappa - \kappa_0) \right]^2 \mathrm{d}v + \alpha_x \int_V w_x \left[\frac{\partial w(z)(\kappa - \kappa_0)}{\partial x} \right]^2 \mathrm{d}v \\ & + \alpha_y \int_V w_y \left[\frac{\partial w(z)(\kappa - \kappa_0)}{\partial y} \right]^2 \mathrm{d}v + \alpha_z \int_V w_x \left[\frac{\partial w(z)(\kappa - \kappa_0)}{\partial z} \right]^2 \mathrm{d}v \end{aligned} \qquad [102]$$

where κ is the susceptibility model, κ_0 is a reference model, and $w(z)$ is the depth-weighting function. The weighting coefficients α_s, α_x, α_y, and α_z are positive numbers controlling the importance of each term, and w_s, w_x, w_y, and w_z are 3-D weighting functions for the smallest and derivative terms, respectively. Equation [102] can be discretized using the model mesh to form a matrix representation:

$$\varphi_m\left(\vec{\kappa}\right) = \left(\vec{\kappa} - \vec{\kappa}_0 \right)^T Z^T W_m^T W_m Z \left(\vec{\kappa} - \vec{\kappa}_0 \right) \qquad [103]$$

where Z is a diagonal matrix containing the discrete weighting function and $W_m^T W_m$ is the model weighting matrix that contains the discretized integrations in eqn [102] and incorporates the coefficients and individual 3-D weighting functions (see Li and Oldenburg, 1996, for details).

Equation [102] is a commonly used model objective function, and its minimization, subject to appropriate data misfit, produces the simplest model that explains the data. The weighting coefficients determine the relative flatness in each of the three spatial directions, whereas the weighting functions determine location wise how close the recovered model is to the reference or how rough or flat the model should be. The combination of the reference model and weighting functions provides a great deal of capability to input descriptive geologic information into the inversion. For example, if geology suggests a rapid change in susceptibility, then a decreased weighting function for derivatives can be put there and the constructed model will exhibit higher gradients provided that this feature does not contradict the magnetic data. The reference model κ_0 may be a general background model that is estimated from previous investigations, or it could be the zero model. Therefore, the reference model and four 3-D weighting functions allow the incorporation into the inversion of additional information other than the magnetic data. The additional information can be from previous knowledge about the susceptibility, from borehole data, from other geophysical inversions, or from the interpreter's qualitative or quantitative understanding about the geologic structure and its relation to susceptibility. When this extra information is incorporated, the inversion derives a model that not only fits the data but, more importantly, also has a likelihood of representing the Earth.

The important component in the model objective function is the depth-weighting function $w(z)$. This function was first developed for inverting surface magnetic data. This function is needed to supply the necessary additional information to construct a geologically plausible susceptibility model because the magnetic data alone are not sufficient. The surface magnetic data are samples of a 2-D function. The data sets at different observation heights are linearly related to each other by the upward continuation operator. 3-D inversions, on the other hand, attempt to construct a 3-D susceptibility model. This difference in dimensionality means that the magnetic data do not have sufficient information to resolve depth. As a result, a minimum-norm solution without depth weighting will only produce models that have nonzero susceptibility concentrated near the surface – the locations closest to the observations.

Mathematically, this difficulty manifests itself as a consistent decrease of the sensitivity $g_{ij}(x, y, z)$ with the increasing depth when the data are located above the surface (ground or

airborne surveys). To counteract this effect, we must supply additional information to the inversion. One approach to do so is to use a depth-weighting function (Li and Oldenburg, 1996). Based on the classical weighted smallest model construction (Parker, 1977, 1994), it can be shown that the square of the depth-weighting function should be equal to the overall decay of the sensitivity with depth. Numerical experiments indicate that the function of the form $(z+z_0)^{-3}$ closely approximates the kernel's decay directly under the observation point, given a correctly chosen value of z_0. This is consistent with the fact that, to first order, a small cubic-shaped cell acts like a dipole source whose magnetic field decays by inverse distance cubed. The value of z_0 can be obtained by matching the function $(z+z_0)^{-3}$ with the kernel function beneath the observation point. Thus, a reasonable candidate for the depth-weighting function is given by

$$w(z) = \frac{1}{(z+z_0)^{3/2}} \qquad [104]$$

where z_0 is a constant to be determined by matching $w^2(z)$ with the change of sensitivity with depth through a least squares fitting. The depth weighting of this form essentially gives all cells at different depths in the model equal probability to become nonzero and represents a type of generic prior information that is not specific to each site.

For a general distribution of observation locations such as highly variable elevation or borehole data, a more general distance weighting or RMS sensitivity weighting (Li and Oldenburg, 2000a) is more appropriate. We will discuss these in a later section.

The third component is the bound constraints imposed on the magnetic susceptibility. For a majority of exploration problems, the magnetic susceptibility has a well-defined range. For example, the susceptibility in most mineral exploration is positive since we are dealing with paramagnetic materials, so the lower bound is zero (positivity). There may also be a known upper bound that should be incorporated into the inversion. Similarly, salt bodies have a negative susceptibility since they are diamagnetic, so the upper bound is zero (negativity). Correspondingly, a lower bound can also be included additionally.

Combining these three components and using the Tikhonov regularization formalism, the problem of inverting for magnetic susceptibility can be stated as a constrained minimization problem:

$$\begin{aligned} &\text{minimize}\, \varphi_d + \beta\varphi_m \\ &\text{subject to}\, b_l \le \kappa \le b_u \end{aligned} \qquad [105]$$

where β is the regularization parameter and b_l and b_u are, respectively, the lower and upper susceptibility bounds.

We now illustrate the algorithm using a single dipping slab model and associated total-field anomaly shown in Figure 34. The slab has a susceptibility of 0.005 SI, and the inducing field has a strength of 50 000 nT and a direction of $(I, D) = (65°, -25°)$. The data have been contaminated by a 2 nT Gaussian noise. Figure 35 shows the inverted susceptibility and corresponding predicted data. It has a maximum value of 0.07, and we have displayed it as a volume-rendered image with a cutoff value of 0.0155.

11.11.7.2 Joint Inversion of Borehole and Surface Data

A general data set containing both surface and borehole data can be inverted using the same methodology described in Section 11.11.7.1. However, the depth-weighting method is no longer applicable. Borehole data are sparsely distributed in 3-D space, and they often surround the volume in which the susceptibility model is to be constructed. In such cases, the sensitivity decays horizontally, away from the boreholes. When combined with surface data, this decay becomes even more complex. As a result, there is not a single dominant direction of sensitivity change as there was in the inversion of surface data alone.

The principle of the depth-weighting function, however, is to weight preferentially the cells of low sensitivity so that all cells have approximately equal probability of deviating from the reference model, and that principle can still be adopted. For a data set containing borehole measurements, we can use a 3-D weighting function that preferentially weights the cells of low sensitivity irrespective of its location:

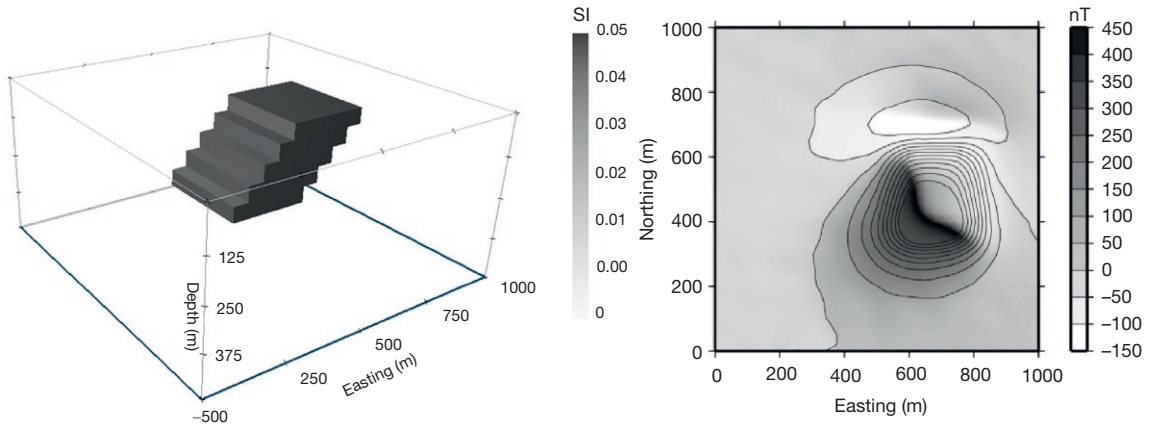


Figure 34 A synthetic example consisting of a susceptible dipping slab buried in a nonmagnetic host. The dipping body has a susceptibility of 0.05 SI. Total-field anomaly data shown on the right. The inducing field has the parameters $(I, D) = (65°, -25°)$.

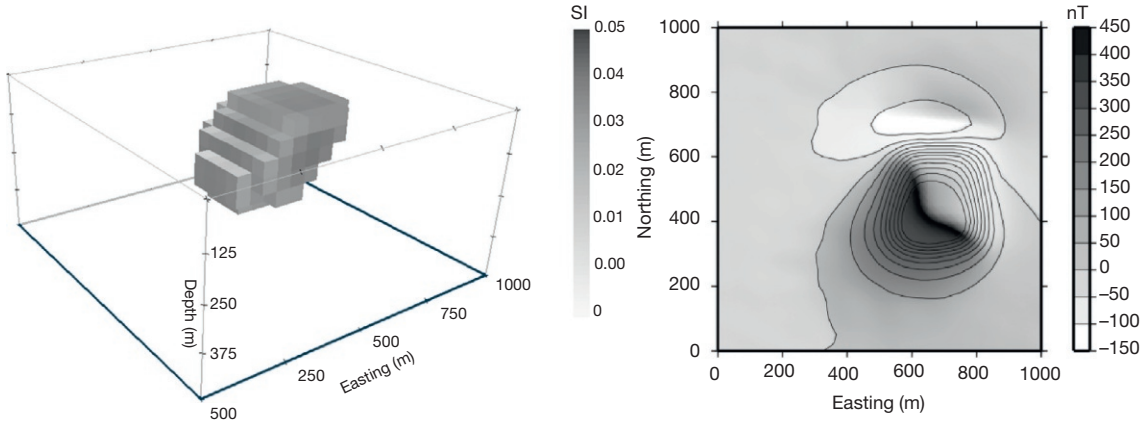


Figure 35 Recovered magnetic susceptibility from the data shown in **Figure 34(b)** and the corresponding predicted data.

$$\varphi_m(\kappa) = \alpha_s \int_V w_s \left[w\left(\vec{r}\right)(\kappa - \kappa_0) \right]^2 \mathrm{d}v + \alpha_x \int_V w_x \left[\frac{\partial w\left(\vec{r}\right)(\kappa - \kappa_0)}{\partial x} \right]^2 \mathrm{d}v$$

$$+ \alpha_y \int_V w_y \left[\frac{\partial w\left(\vec{r}\right)(\kappa - \kappa_0)}{\partial y} \right]^2 \mathrm{d}v + \alpha_z \int_V w_x \left[\frac{\partial w\left(\vec{r}\right)(\kappa - \kappa_0)}{\partial z} \right]^2 \mathrm{d}v \qquad [106]$$

A general weighting function $w\left(\vec{r}\right)$ (with $\vec{r}$ denoting the location in the model domain) is needed in place of the depth weighting because the observations are located in 3-D space and the sensitivity decay is not dominated in any particular direction as in the case of surface data. There are two approaches to generating a 3-D weighting function (Li and Oldenburg, 1996, 2000a). The first is based on the sensitivity matrix and extracts the general decay by the overall sensitivity of the entire data set to a particular cell in the model, and the second is a generalization of the method used in surface data inversion.

11.11.7.2.1 Sensitivity-based weighting function

Assuming that the problem has been discretized, Li and Oldenburg (2000a,b) defined an RMS sensitivity, S_j, as the measure of the sensitivity of the entire data set to the jth cell:

$$S_j = \left(\sum_{i=1}^{N} g_{ij}^2 \right)^{1/2}, \quad j = 1, \ldots, M \qquad [107]$$

where g_{ij} is the sensitivity of ith datum with respect to the jth cell, N is the number of data, and M is the number of cells in the model. Therefore, we can use the square root of the RMS sensitivity as a general weighting function when inverting borehole magnetic data:

$$w_j = \left(\sum_{i=1}^{N} g_{ij}^2 \right)^{1/4}, \quad j = 1, \ldots, M \qquad [108]$$

When there are only surface data at a constant elevation, w_j agrees with the depth weighting in the preceding subsection, and the inversion using this general weighting function produces consistent results. For general data sets such as borehole data sets, this weighting function captures the general decay of the magnetic kernel with distance from the nearest data locations. A similar form of weighting function has been used by Wang (1995) in seismic refraction tomography to deal with the uneven distribution of ray paths. The same weighting function was later used by Portniaguine and Zhdanov (2002) in their magnetic inversion algorithm.

This weighting function has a minor drawback. Since the magnetic kernels are direction-dependent, the resulting weighting function will vary with the azimuth of a cell relative to a borehole. This azimuthal variation is clearly visible when only single-component data are available; however, it is eliminated when three-component data are used.

11.11.7.2.2 Distance-based weighting function

To overcome this drawback of the RMS sensitivity-based weighting, we can use a weighting function based purely on the distance between the cell and observation locations:

$$w_j = \left(\sum_{i=1}^{N} \int_{\Delta V_j} \frac{dv}{\left(\left| \vec{r} - \vec{r}\,' \right| + R_0 \right)^3} \right)^{1/4}, \quad j = 1, \ldots, M \qquad [109]$$

where R_0 is a constant similar to z_0 in eqn [104]. This weighting function essentially combines the distance characteristics of the original depth weighting for surface data with the generality of the weighting function based upon RMS sensitivity. It is independent of the sensitivity calculation and is not affected by the direction of measured anomaly component.

We now illustrate the method with a synthetic example. The true susceptibility model consists of two prisms buried at different depths in a nonsusceptible half-space. The inducing field is assumed to be at $I = 65°$ and $D = 25°$. The total-field anomaly was calculated on the surface at an interval of 25 m along east–west lines spaced 100 m apart, resulting in a total of 175 data. Three-component borehole data are calculated in three holes. There are 16 observation locations spaced 25 m vertically in each hole, and the total number of data in three holes is 144. Gaussian noise having a standard deviation of 2 nT has been added to all data, and the resulting simulated observations are shown in **Figure 36**. The weighting functions corresponding to the surface data, borehole data, and the joint data set are

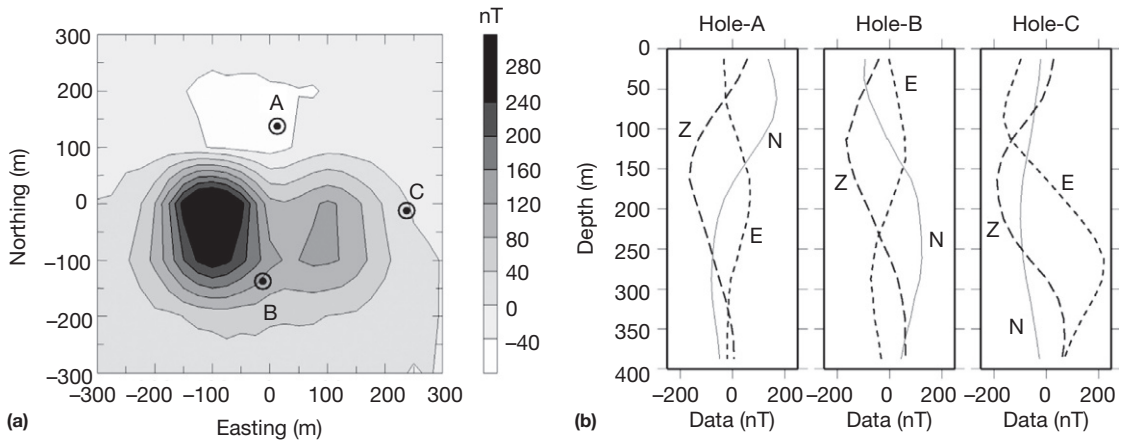


Figure 36 The left panel is the total-field anomaly on the surface produced by a model consisting of two vertical prisms in an inducing field in the direction $I = 65°$ and $D = 25°$. The locations of three vertical boreholes are indicated by A, B, and C. The three panels on the right are the three-component anomalies in these boreholes. Different components are identified by the labels beside the curves. Reproduced from Li Y and Oldenburg DW (2000a) Joint inversion of surface and three-component borehole magnetic data. *Geophysics* 65: 540–552; Li Y and Oldenburg DW (2000b) Reduction to the pole using equivalent sources. *60th Annual International Meeting*, *SEG*, *Expanded Abstracts*, pp. 386–389.

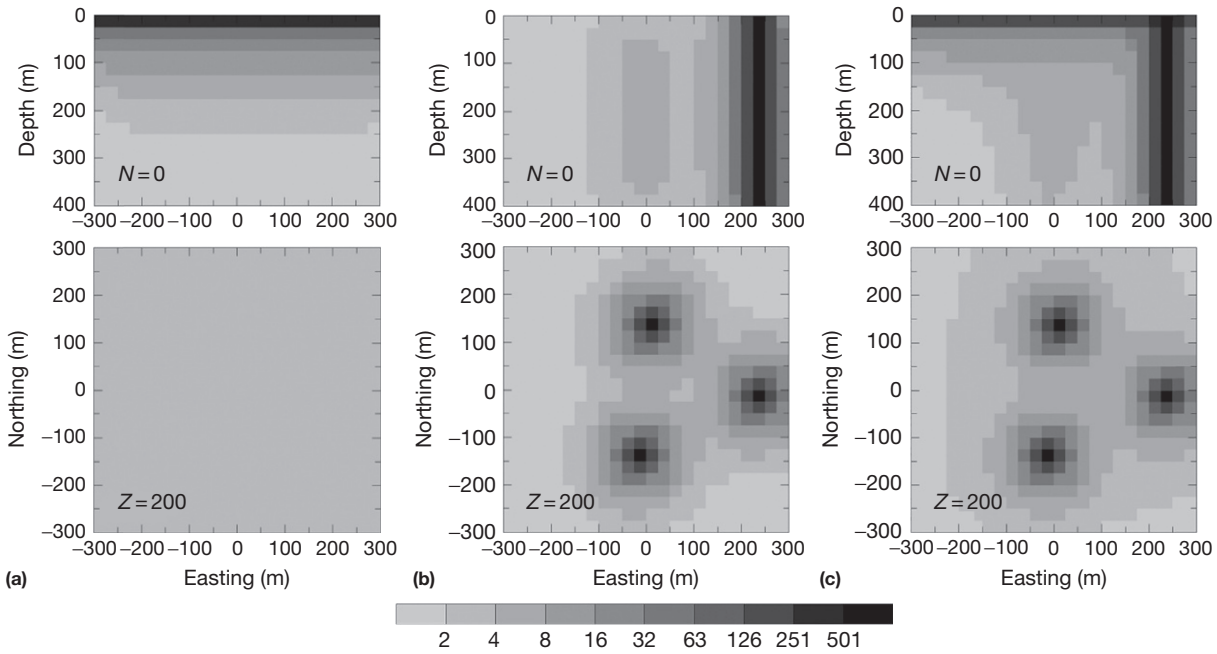


Figure 37 The general weighting function based on the RMS sensitivities of different data sets. The squares of the weighting functions are shown here, and the dimensionless value is multiplied by 1000 for display purposes. (a) One cross section and one plan section of the weighting function for the surface data shown in **Figure 36**. (b) The borehole data shown in **Figure 36**. The weighting is highest near the boreholes. (c) The joint data set, consisting of surface and borehole data. Reproduced from Li Y and Oldenburg DW (2000a) Joint inversion of surface and three-component borehole magnetic data. *Geophysics* 65: 540–552; Li Y and Oldenburg DW (2000b) Reduction to the pole using equivalent sources. *60th Annual International Meeting*, *SEG*, *Expanded Abstracts*, pp. 386–389.

shown in Figure 37. We observed that the weighting for the surface data is dominantly decaying with the depth, the weighting for the borehole data is primarily a function of the horizontal location and peaks near the borehole, and the weighting for the joint data sets captures the characteristics of the previously mentioned two. Using the weighting function in Figure 37(c), the result of inverting the joint data set is shown in Figure 38. We observe that there is good definition of both magnetic prisms, whose true positions are outlined by the white boxes.

11.11.7.3 Computational Aspects

Successful application of the previously described algorithm hinges upon the ability to carry out the inversions efficiently in practice. The two computational challenges we face are the

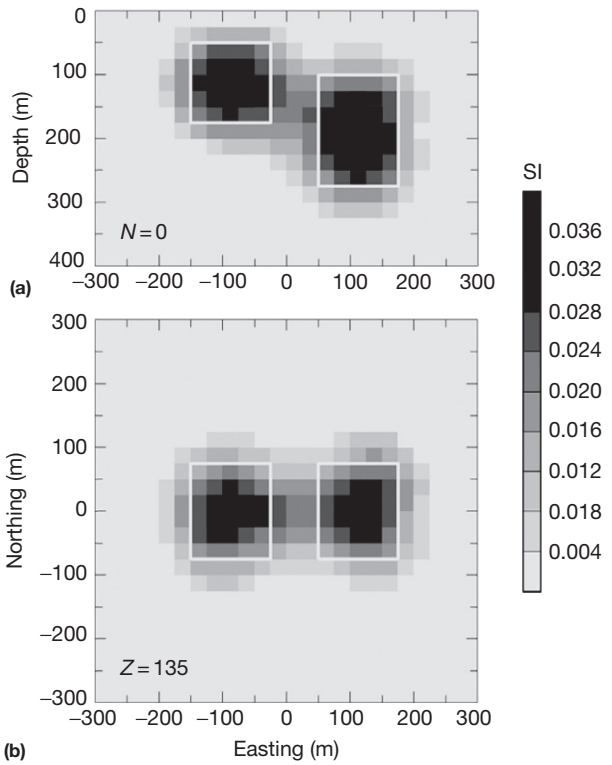


Figure 38 The susceptibility model recovered from the joint inversion of total-field surface data and three-component borehole data. The sensitivity-based weighting function is used in this inversion. This model clearly defines both prisms and illustrates the improvement achieved by joint inversion of surface and borehole data. Reproduced from Li Y and Oldenburg DW (2000a) Joint inversion of surface and three-component borehole magnetic data. *Geophysics* 65: 540–552; Li Y and Oldenburg DW (2000b) Reduction to the pole using equivalent sources. *60th Annual International Meeting*, *SEG*, *Expanded Abstracts*, pp. 386–389.

imposition of the bound constraints and the solution of the optimization problem for large-scale data sets. We address these issues in the succeeding text.

11.11.7.3.1 Bound constraints through logarithmic barrier method

Bound constraints play an important role in potential-field inversions because of the lack of sufficient information in these data. The need for bounds is particularly important in magnetic inversion because of the complexity introduced by the magnetization direction and the projection of anomaly onto a given direction, that is, the inducing field direction. As a result, the inversion is solved as an inequality-constrained optimization (eqn [105]).

There are many different approaches to impose the constraints. One is to introduce a nonlinear mapping function and an auxiliary variable so the new variable can take on unrestricted values, while the function has a range within the bounds (Li and Oldenburg, 1996). Another approach is to work with the logarithm of the susceptibility as the model. Alternatively, the primal logarithmic barrier method is easy to implement and computationally efficient for this purpose.

The logarithmic barrier method was originally developed for solving linear and quadratic programming problems with inequality constraints (e.g., Gill et al., 1991; Wright, 1997). Li and Oldenburg (2003) introduced this approach into magnetic inversion, whereby the bound constraints are imposed by adding logarithmic terms in the total objective function of the Tikhonov regularization:

$$\varphi = \varphi_d + \beta\varphi_m - 2\lambda\sum_{j=1}^{M}\left(\ln\frac{k_j - b_l}{b_u - b_l} + \ln\frac{b_u - \kappa_j}{b_u - b_l}\right) \qquad [110]$$

where the third term is the logarithmic barrier term and λ is the barrier parameter. As the name suggests, the logarithmic barrier term forms a barrier along the lower and upper bounds, because as the susceptibility approaches either lower or upper bound, the logarithm approaches $-\infty$ and the barrier term approaches ∞. Consequently, it prevents the minimization from producing susceptibility values outside the bounds. The method solves a sequence of nonlinear minimizations with decreasing barrier parameter λ, and, as $\lambda \to 0$, the sequence of solutions to the minimization problem in eqn [110] approaches that of the problem in eqn [105].

The minimization process starts with a large value of λ and an initial model $\vec{\kappa}$ whose elements are all within bounds. It then finds the solution iteratively with the barrier parameter λ being decreased at each iteration. Instead of carrying out the full minimization in eqn [110] at each iteration, we take one Newton step for each value of λ and adjust the step length so that the updated model remains within bounds (e.g., Gill et al., 1991). The step length is also used to determine the decreased value of the barrier parameter λ. For a fixed regularization parameter, the model perturbation from a Newton step at the nth iteration is given by

$$\begin{aligned}&\left(G^T W_d^T W_d G + \beta Z^T W_m^T W_m Z + \lambda^{(n)} X_l^{-2} + \lambda^{(n)} X_u^{-2}\right)\Delta\vec{\kappa} \\ &= G^T W_d^T W_d\left(\vec{d}^{\,\mathrm{obs}} - \vec{d}^{\,(n)}\right) + \beta Z^T W_m^T W_m Z\left(\vec{\kappa}_0 - \vec{\kappa}^{\,(n)}\right) \\ &\quad -\lambda^{(n)}\left(X_l^{-1} + X_u^{-1}\right)\vec{e}\end{aligned} \qquad [111]$$

where $X_l = \mathrm{diag}(\ldots, \kappa_j^{(n)} - b_l, \ldots)$, $X_u = \mathrm{diag}(\ldots, b_u - \kappa_j^{(n)}, \ldots)$, $\vec{e} = (1, \ldots, 1, \ldots, 1)^T$, and $\vec{\kappa}(n)$ and $\vec{d}(n)$ are, respectively, the current model and predicted data at the nth iteration. The model is updated by

$$\vec{\kappa}(n+1) = \vec{\kappa}(n) + \gamma\alpha\Delta\vec{\kappa} \qquad [112]$$

where α is the maximum step length that will bring the updated model to the bounds and $\gamma < 1.0$ is a user-specified constant to ensure that the updated model remains within bounds. The iteration continues with a reduced barrier parameter $\lambda^{(n+1)}$ until the barrier term is much smaller than the data misfit and model objective function.

As in all inversions, this process is repeated over a range of the regularization parameter β to determine the value that yields the expected data misfit. Different strategies and ad hoc approaches can be used to speed up this search process.

11.11.7.3.2 Wavelet compression-based acceleration

We now proceed to the last component of a practical inversion algorithm, namely, the efficient solution of eqn [111]. We note

that for large-scale problems, it is unrealistic to use a direct solver by first forming and storing the matrix on the left-hand side of eqn [111]. For example, if we have a $100 \times 100 \times 50$ model mesh, and 10 000 data points, which is a moderate-sized problem, just storing the sensitivity matrix G in single precision will require more than 19 GB of memory. For this reason, it is now standard to solve eqn [111] by an iterative approach. Given that the matrix is real and symmetrical, the CG method (Hestenes and Stiefel, 1952; Nocedal and Wright, 1999) is ideal. This approach was used by Pilkington (1997) in magnetic inversion and later by Li and Oldenburg (2003). Such an approach is now standard.

When using the CG method, one only needs to calculate the result of multiplying the matrix on the left-hand side of eqn [111] with a vector without explicitly having that matrix. Of the four components, the last three, $Z^T W_m^T W_m Z$, X_l^{-2}, and X_u^{-2}, are extremely sparse, and they do not require significant memory to storage or CPU time to multiply with a vector. The majority of the computation lies in the multiplication of $G^T W_d^T W_d G$ with a vector. As discussed previously, the challenges are in storing the matrix and carrying out the multiplication.

The compression of the sensitivity matrix G (Li and Oldenburg, 2003) using wavelet transforms based on the compactly supported orthonormal wavelets offers an effective solution to both challenges. Wavelets (e.g., Daubechies, 1992; Mallat, 1989; Meyer, 1993) are recursively defined by the dilation and translation of a mother wavelet. A class of orthonormal and compactly supported wavelets are constructed by solving a two-scale difference equation (Daubechies, 1992). These wavelets have several important properties: (1) They are orthogonal to the dyadic translation and dilation of the original version, (2) they have compact support and are localized in both the space domain and the frequency domain, and (3) they are constructed orthogonal to low-order polynomials.

The discrete wavelet transform using these orthonormal, compactly supported wavelets represents a discrete function by a set of wavelet coefficients. An efficient algorithm for computing such an orthonormal wavelet transform is the well-known pyramid algorithm first outlined by Mallat (1989). The discrete wavelet transform is a linear and orthonormal transformation, and it can be symbolically represented by a matrix that acts upon a vector representing the discretized function. Let $\mathbf{W}$ be the matrix representing the fast wavelet transform and $\mathbf{v}$ be the vector; then, the wavelet transform $\widetilde{v}$ is given by

$$\widetilde{v} = \mathbf{W}\mathbf{v} \qquad [113]$$

The transform $\mathbf{W}$ is orthonormal and its inverse is given by its transpose, $\mathbf{W}^{-1} = \mathbf{W}^T$. This relation allows the reconstruction of a function from its wavelet coefficients:

$$\mathbf{v} = \mathbf{W}^T \widetilde{v} \qquad [114]$$

A large number of the wavelet coefficients are either zero or very close to zero. Winnowing the coefficients whose magnitudes are below a certain level still allows the reconstruction of the original function with a high degree of accuracy. Thus, a function can have a sparse representation in the wavelet domain. This results first from the localization property of the wavelets.

The property of wavelet sparse representation is illustrated in **Figure 39**. The wavelet transform (using Daubechies-4 wavelet) of a magnetic anomaly profile. There are 128 data points in the profile. The wavelet coefficients are shown as

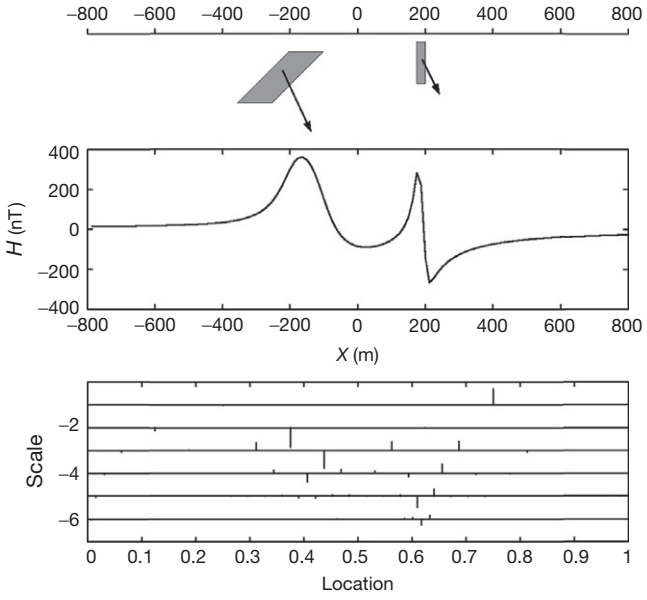


Figure 39 Fast wavelet transform of a magnetic anomaly profile. The 2-D distribution of susceptibilities is shown in top panel. The magnetic data are in the middle panel. The wavelet coefficients are shown as vertical lines with the length proportional to the amplitude. They are plotted as a function of location and scale. Reproduced from Li Y and Oldenburg DW (2003) Fast inversion of large-scale magnetic data using wavelet transforms and logarithmic barrier method. *Geophysical Journal International* 152: 251–265.

vertical lines with the length proportional to the amplitude at their locations and scales. Notice that most of the wavelet coefficients are small. Discarding the coefficients that are below a threshold relative to the largest coefficient produces a sparse representation of the function in the wavelet domain. The original function can be reconstructed with a high degree of accuracy from the sparse representation.

In the cases of multidimensional functions, one way of accomplishing the multidimensional wavelet transform is by applying the 1-D wavelet transforms independently to each dimension (e.g., Daubechies, 1992). This is analogous to the FFT in multidimensions. Such a multidimensional wavelet transform amounts to using 3-D wavelets formed by the tensor products of 1-D wavelets.

The rows of the sensitivity matrix are defined over the 3-D domain of the model κ and can be treated as 3-D images. The application of the 3-D wavelet transform can therefore produce a sparse representation of them. Let $\mathbf{W}$ be the 3-D wavelet transform and $\widetilde{\mathbf{G}}$ be the matrix whose rows are the wavelet transforms of corresponding rows of $\mathbf{G}$. Then, the transformed sensitivity matrix $\widetilde{\mathbf{G}}$ is given by

$$\widetilde{\mathbf{G}} = \mathbf{G}\mathbf{W}^T \qquad [115]$$

We also apply the identical wavelet transform to the model vector κ to produce the transformed model $\widetilde{\kappa}$. Let

$$\widetilde{\kappa} = \mathbf{W}\kappa \qquad [116]$$

Multiplying eqns [115] and [116] and applying $\mathbf{W}^T\mathbf{W} = \mathbf{I}$ yields

$$\widetilde{\mathbf{G}}\widetilde{\kappa} = \mathbf{d} \qquad [117]$$

Therefore, the forward modeling is accomplished by the multiplication of the transformed sensitivity and the transformed model in the wavelet domain. Equation [117] allows the application of the sensitivity matrix to a general model vector during the CG solution. We also require the application of $\mathbf{G}^T$ to a data vector $\vec{v}_d$ to produce a 'model' vector $\mathbf{v}_{\kappa'}$ which is given by

$$\mathbf{W}^T\widetilde{\mathbf{G}}^T\vec{v}_d = \vec{v}_\kappa \qquad [118]$$

Thus, the application of $\mathbf{G}^T$ to a vector is accomplished by the multiplication of the transformed sensitivity with the data followed by an inverse wavelet transform.

The transformation to the wavelet domain itself does not decrease the required storage nor does it increase the speed of calculation. It is the properties of functions in the wavelet domain, and the processing based upon these properties, that make the fast algorithm possible. Thus, we can apply thresholding to the transformed sensitivity matrix $\widetilde{G}$ to generate a sparse matrix $\widetilde{G}^s$ by retaining only significant wavelet transform coefficients.

During the course of an inversion, we need only to store the sparse matrix $\widetilde{G}^s$, which yields the desired reduction in the memory requirement. In addition, the matrix $\widetilde{G}^s$ is substituted for $\widetilde{G}$ in eqns [117] and [118] so that the multiplication of G or G^T with a vector is carried out by sparse multiplication in the wavelet domain and the number of operations is reduced in proportion to the sparseness of ν. This yields the desired reduction in CPU time.

A typical compression ratio that can be achieved in practice can range from 10 to 100 using the previously mentioned method. That means the storage required by the wavelet compression is 5–50 times smaller, assuming we use a row major format to store the sparse transformed matrix, while the computational time for solving the inverse problem is ~10–100 times shorter. These represent significant savings in computational time, enabling us to solve much larger problems with the same amount of computational power.

Davis and Li (2013) had extended this approach to adaptive octree mesh by utilizing the Hilbert space filling curve to order the cells in a semistructured octree mesh and then applying a 1-D wavelet transform to each row of the reordered sensitivities. This approach is able to achieve a much higher compression ratio and the combined compression ratio using the octree and wavelet transform can reach as high as 1000.

11.11.7.4 A Field Example

The successful application of 3-D magnetic inversion by Falconbridge at the Raglan deposit in northern Quebec exemplifies the utility of modern 3-D magnetic inversion in mineral exploration (Oldenburg et al., 1998; Watts, 1997). The Raglan mine operated by Falconbridge is located in northern Quebec and the nickel mineralization is hosted in ultramafic flow units. Little surface expression of geology is available, and exploration has relied primarily on geophysics and magnetics in particular. Total-field magnetic data in the area are typified by seemingly isolated magnetic highs interconnected by arc-like low-intensity anomalies. The top panel in **Figure 7** shows an example area of 4 km by 4 km. Two regions of high magnetic field values are observed and they correspond to highly magnetic ultramafic outcrops, which contain economic-grade ores. The geologic question was whether the outcrops were associated with a single flow unit or isolated bodies. The answer has important implications since the former scenario means there is great potential for extending the reserve beyond the known surface deposit. Generalized 3-D inversion was then applied (Oldenburg et al., 1998; Watts, 1997).

For the 3-D inversion, the volume of the Earth below the data area is represented with 40 by 40 by 10 cubic cells. The dimension of the cells was 100 m. The model objective function is designed to make the final susceptibility distribution smooth in all three spatial directions and close to a zero background. Positivity constraints and a depth-weighting function were included in the inversion. The result of inversion indicates the presence of a continuous zone of highly magnetic material that extends between the two outcrops. The zone is shown as a volume-rendered image (**Figure 7**, bottom panel) generated by Falconbridge with a threshold level of 0.04. Since the only known magnetic rock in the area is ultramafic flow, it is reasonable to equate the presence of highly magnetic material found through inversion to the existence of ultramafic flows. It was this image (**Figure 40**, bottom panel) that convinced the project geologist that the apparently isolated outcrops were in fact connected at depth. The deep 1100 m hole located based on this image intersected the magnetic source at 650 m. As a bonus, a 10 m mineralized section (subore grade, ~1% Ni) was intersected within the 350 m thick

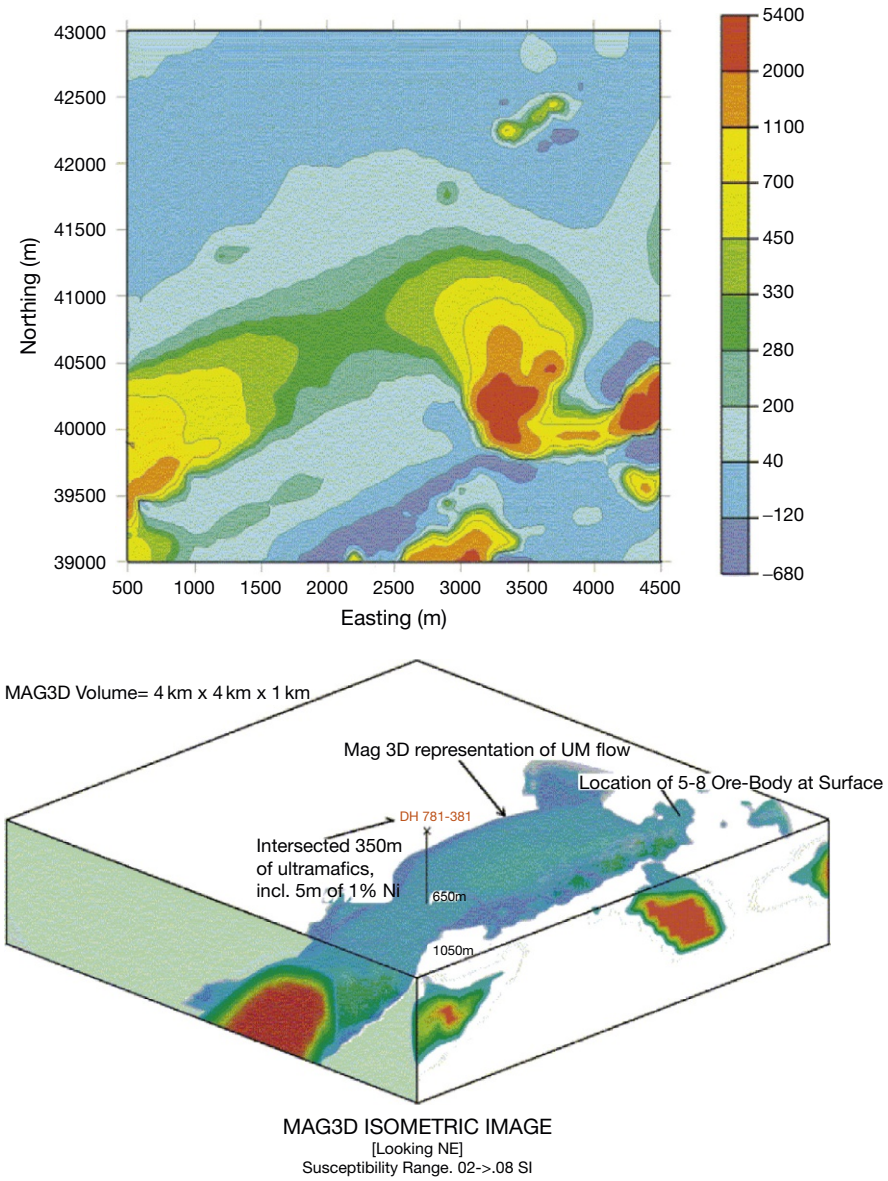


Figure 40 The top panel shows the total-field aeromagnetic data. The inducing field has $I = 83^\circ$ and $D = -32^\circ$. Data are contoured in nanotesla. The lower panel shows the 3-D susceptibility model recovered from magnetic inversion at Raglan deposit. This is a volume-rendered image of the inverted susceptibility model, and the displayed surface provides a representation of the ultramafic flow. Indicated in the east–west section and in the volume-rendered representation is the intersecting drill hole that was located based upon the inversion result. Reproduced from Watts A (1997) Exploration for nickel in the 90s, or 'Till depth us do part.' *Exploration 97*.

intersection of magnetic ultramafics. Subsequently, 3-D inversion has been used extensively in this region and new geologic horizon and economic reserves have been found as a direct consequence of applying 3-D magnetic inversion (Watts, personal communication).

11.11.8 Magnetization Inversion

The key assumption in magnetic susceptibility inversion discussed in the preceding section is that the magnetization is induced only and the self-demagnetization effect is negligible, so that the magnetization direction is known and given by the inducing field direction. The inversion recovers the susceptibility that directly defines the magnitude of the magnetization. However, there are well-documented cases in which such an assumption is inadequate due to the presence of remanent magnetization and strong self-demagnetization effects.

The presence of remanent magnetization is widespread and can occur in rock units having either small or large magnetic susceptibility. The presence of strong magnetic susceptibility that leads to self-demagnetization also occurs in many exploration and environmental problems. The total magnetization is the vector sum of an induced component and the remanent component in weakly magnetic source bodies. The magnetization direction in the presence of high susceptibility can be

highly variable due to the geometry dependence of the self-demagnetization effect. In both cases, the direction of the magnetization is unknown and can be significantly different from that of the inducing field. Therefore, the inversion ultimately needs to determine both the direction and magnitude of the magnetization. Consequently, the approach in the preceding section is not directly applicable. For instance, simulations by Shearer (2005) have indicated that the algorithm by Li and Oldenburg (1996) yields erroneous results when the error in the specified magnetization direction exceeds 15°. This difficulty can severely limit the application of that class of algorithms.

There are a number of papers published about the self-magnetization effect (e.g., Lelievre and Oldenburg, 2006; Wallace, 2007), but the majority of the effort has been on remanent magnetization. Several different approaches have been developed to tackle this problem. The first is to estimate the total magnetization direction and then supply the direction to the magnetic susceptibility inversion. The second is the amplitude inversion that skirts the direction issue and inverts for the magnitude of total magnetization (Li et al., 2010; Shear, 2005). This approach is more effective when the magnetization direction is highly variable due to factors such as structural changes and multiple sources formed over different geologic times. Another approach in the same category is to invert the normalized source strength (Beiki et al., 2012; Pilkington and Beiki, 2013) derived from the eigenvalue decomposition of the magnetic gradient tensor. The third is to invert for the magnetization vector itself (Ellis et al., 2012; Lelièvre and Oldenburg, 2009; Li and Sun, 2013; Liu et al., 2013; Wang et al., 2004). This approach is generally applicable but may require site-specific constraints to offset the extra non-uniqueness introduced by the additional model parameter sets to be recovered. A similar approach based on parametric inversion is to invert for the magnetization vectors within causative bodies together with their geometric shapes (e.g., Foss and McKenzie, 2011; Mueller et al., 1997; Silva and Hohmann, 1981).

11.11.8.1 Direction Estimation

For isolated anomalies, one may estimate the magnetization direction from the magnetic data directly. This direction can then be supplied to the susceptibility inversion described in **Section 11.11.7** to recover a susceptibility distribution in the subsurface. The recovered susceptibility does not correspond to intrinsic susceptibility, but represents the combined source strength due to induced and remanent magnetization components. It can be conveniently referred to as the effective susceptibility, which is the ratio of the total magnetization over the inducing field strength.

Magnetization direction estimation is a well-studied topic in magnetic exploration. For example, Zietz and Andreasen (1967) examined the relationship between position and intensity of the maximum and minimum produced by a simple causative body. Roest and Pilkington (1993) correlated the amplitude of the 2-D total gradient of the magnetic field and the absolute value of the horizontal gradient of the pseudo-gravity produced by 2-D sources. Lourenco and Morrison (1973) and Phillips (2005) developed a method based upon the integral relationships of magnetic moments derived by Helbig (1963). More recently, Haney and Li (2002) developed a wavelet-based method for determining magnetization direction in 2-D data sets. Dannemiller and Li (2006) introduced an improvement to Roest and Pilkington (1993) cross correlation method and extended it to 3-D cases. Gerovska et al. (2009) also developed a similar method.

Here, we present the Helbig moment, wavelet, and two cross correlation methods. The first two directly explore the relation between the anomaly and the magnetization direction and compute the magnetization direction from the data, whereas the third estimates the magnetization direction by utilizing the symmetry property of the reduced-to-the-pole (RTP) field. In both categories, we assume the magnetization direction does not vary drastically for sources within the volume under examination.

11.11.8.1.1 Helbig moment method

The Helbig method (Helbig, 1963; Lourenco and Morrison, 1973; Phillips, 2005) calculates the magnetization vector components m_x, m_y, and m_z of a compact magnetic source from the first moments of the vector components B_x, B_y, and B_z of the anomalous magnetic field:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xB_z(x,y)\mathrm{d}x\mathrm{d}y = -2\pi m_x$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yB_z(x,y)\mathrm{d}x\mathrm{d}y = -2\pi m_y \qquad [119]$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xB_x(x,y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yB_y(x,y)\mathrm{d}x\mathrm{d}y = -2\pi m_z$$

In the equations mentioned earlier, the three components of the total magnetization vector (induced plus remanent) are assumed to be constant within the source body.

Two scenarios arise in practical applications. First, we usually have only the total-field anomaly data, and one needs to calculate the three components B_x, B_y, and B_z from the total-field anomaly using wave number-domain operators (e.g., Blakely, 1996; Pedersen, 1978; Schmidt and Clark, 1997, 1998). Difficulties may arise when the data are acquired in low magnetic latitudes since the conversion requires an HRTP operation. However, as vector magnetic surveys are becoming available (e.g., Dransfield et al., 2003), the observed three-component data can be used directly in the Helbig method. Second, since the method is applied successively over finite data windows and since the integrals of B_x, B_y, and B_z over the infinite plane of integration (their mean values) are equal to zero, care must be exercised to satisfy these conditions within the data window over which the calculations are made. This is achieved by removing planar surfaces from each window (Phillips, 2005).

To overcome the previously mentioned difficulty associated with the unknown DC component, Nabighian (2006) proposed extending the Helbig method to first-order derivatives of the magnetic components instead of the components themselves. These derivatives can either be calculated or be measured directly with tensor magnetic gradiometer systems.

A simple integration by parts of expressions [119] and taking into account that the magnetic components decay as $1/r^3$ lead immediately to

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2\frac{\partial B_z(x,y)}{\partial x}\mathrm{d}x\mathrm{d}y = 4\pi m_x\\
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^2\frac{\partial B_z(x,y)}{\partial y}\mathrm{d}x\mathrm{d}y = 4\pi m_y\\
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2\frac{\partial B_x(x,y)}{\partial x}\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^2\frac{\partial B_y(x,y)}{\partial y}\mathrm{d}x\mathrm{d}y = 4\pi m_z\\
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\frac{\partial B_z(x,y)}{\partial y}\mathrm{d}x\mathrm{d}y = 2\pi m_x\\
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\frac{\partial B_z(x,y)}{\partial x}\mathrm{d}x\mathrm{d}y = 2\pi m_y\\
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\frac{\partial B_x(x,y)}{\partial y}\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\frac{\partial B_y(x,y)}{\partial x}\mathrm{d}x\mathrm{d}y = 2\pi m_z
\end{aligned} \qquad [120]$$

The fact that one uses the component derivatives instead of the component themselves implies that one need not remove anymore a planar surface from data to satisfy the requirement of zero mean in each window, thus yielding superior results. A complete treatment of the extended Helbig method can be found in Phillips et al. (2007) where it is also combined with the Euler deconvolution for a more accurate source location and depth determination.

11.11.8.1.2 Wavelet multiscale edge method

Haney and Li (2002) developed a method for estimating the magnetization direction in 2-D using multiscale edges of a magnetic anomaly derived through a CWT. The multiscale edges correspond to the trajectories of the extrema of the wavelet transform of the anomaly profile, and their positions in the $x-z$ plane are dependent upon the inclination of magnetization. Tracking the multiscale edges allows one to determine the inclination of the magnetization in 2-D sources.

Given a profile of magnetic data observed at a height z_0 in an inducing magnetic field with inclination I, a CWT can be performed using a set of natural wavelets that is equivalent to the magnetic field produced by a line dipole in a particular direction (Hornby et al., 1999). When the dipole has an inclination of $-I$, the corresponding CWT depends only upon the magnetization direction.

The magnitude of the multiscale edges depends on the source geometry, whereas the location depends primarily on the intrinsic properties of the source, that is, the magnetization direction I_m. For example, for a line of dipoles, there are four trajectories and they are given by

$$s = x\cot\left(\frac{I_m \pm \pi/2}{3}\right),\ x\cot\left(\frac{I_m \pm 3\pi/4}{3}\right) \qquad [121]$$

where s is the dilation factor that corresponds to the ratio of upward continuation height over the original observation height and x is the location along the magnetic profile. The dipole line is assumed to be directly below $x=0$. The problem of estimating the direction of magnetization therefore becomes one of tracking the trajectories of the multiscale edges and calculating I_m by regression according to eqn [121].

The approach mentioned earlier can be applied to 3-D data sets. In such cases, we can carry out integration of the data in each of the two horizontal directions to synthesize two profiles. Integrating in the east–west direction simulates a profile along a north–south traverse above a 2-D source that strikes in the east–west direction. The translation-invariant property of magnetic anomalies means that integrating the anomalous field is equivalent to adding 3-D bodies along the strike direction to build up a 2-D causative body. The corresponding magnetization is given by the projection of the 3-D magnetization vector in the north–south cross section. Applying the wavelet estimation to this profile produces the apparent inclination of the magnetization within the north–south section. Performing similar operations in the perpendicular direction yields the apparent inclination in the east–west section. The true inclination and declination of magnetization in the original 3-D source can then be reconstructed from these two apparent inclinations in the easting and northing cross sections, respectively.

11.11.8.1.3 Cross correlation methods

The RTP anomaly for a symmetrical vertical body is symmetrical. For dipping bodies, the RTP field is the least asymmetrical. It follows that the vertical derivative of RTP anomaly is therefore also least asymmetrical. It has also been shown that the total gradient (amplitude of the gradient vector in 3-D) of the RTP anomaly is the envelope of the vertical derivative of the anomaly produced under arbitrary inducing field and magnetization directions (Haney et al., 2003). The envelope, by definition, is the most symmetrical form. Utilizing these properties, Dannemiller and Li (2006) developed a method for estimating magnetization by examining the symmetry of various RTP fields. This method was an extension of the 2-D cross correlation method developed by Roest and Pilkington (1993) to 3-D and improves upon the latter by using two quantities that have the same decay rate with distance to the source.

The method searches for the particular magnetization direction that yields the maximum symmetry in the resultant RTP field. The RTP process transforms measured total-field anomalies to the vertical component of the magnetic field that would be generated by a magnetization that has the same magnitude but is oriented in the vertical direction. This transformation can only be performed correctly when the total magnetization direction $\hat{M}_t$ is known. The operation can still be carried out numerically if an incorrect direction is used, but the resultant quantity, $R(\hat{M})$, is not the RTP field and will in general be more asymmetrical than the true RTP field. The problem of finding the correct direction, $\hat{M}_t$, is then one of determining when the resultant RTP field is the least asymmetrical. The symmetry of RTP field can be measured by the cross correlation between the vertical derivative and total gradient of the RTP anomaly that is calculated using an assumed magnetization direction.

The envelope of magnetic anomaly can be readily computed for 2-D problems (Nabighian, 1972), and it can be used as the reference to gauge the degree of symmetry of the vertical gradient of $R(\hat{M})$. In 3-D problems, such an envelope cannot be computed without the knowledge of magnetization direction (Haney et al., 2003). As a result, we do not have a

fixed reference for comparison in 3-D. However, the total gradient and the vertical gradient of $R(\hat{M})$ simultaneously achieve maximum symmetry when a correct magnetization direction is used in the calculation, and their correlation is maximized. The correlation between these two gradients for an arbitrary $R(\hat{M})$ is weaker. Thus, comparing the vertical gradient of $R(\hat{M})$ with its total gradient would enable us to find the correct direction $\hat{M}_t$.

Although there is no analytic proof to show that these two quantities should have maximum correlation for the true RTP field, its practical validity is confirmed through empirical observations in 2-D. Based on this new criterion, the total magnetization direction in 3-D can be estimated by the following procedure. We apply the RTP operator to the observed magnetic anomaly over a range of magnetization directions defined by inclination and declination pairs. After obtaining the RTP field $R(\hat{M})$, we determine the vertical and total gradient maps and compute the zero-lag cross correlation between the two. The assumed total magnetization direction that maximizes the correlation defines the sought magnetization direction.

The estimation process becomes one of searching for the inclination and declination that maximize the cross correlation. Furthermore, we note that the choice of total gradient and vertical gradient is motivated by our desire to formulate a method applicable to dipping bodies and to cases with causative bodies located at different depths. Both quantities decay with depth at the same rate in 3-D, and for a dipolar source, they are both proportional to $1/r^4$. The relative contributions to both quantities from sources at different depths are similar; therefore, the shapes of the total gradient and vertical gradient are affected in a similar manner by source depth. For this reason, the approach works well also in the case of dipping source bodies.

Gerovska et al. (2009) developed a similar estimation method based on the cross correlation between the RTP field and the amplitude of the anomalous magnetic vector. This method is based on the fact that the amplitude has a weak dependence on the magnetization direction and can be used to gauge the symmetry of the RTP field.

The key operation of these methods is the RTP process. Consequently, we will encounter difficulties at low magnetic latitudes as before. Fortunately, several stable RTP methods are available for low latitudes (e.g., Hansen and Pawlowski, 1989; Li and Oldenburg, 2001; Mendonca and Silva, 1993). There is also a trade-off between the accuracy of estimated inclination and the accuracy of estimated declination. The former improves as it approaches −90° or 90°, but the accuracy of the corresponding declination decreases. However, this does not pose a major problem since the influence of the declination becomes less important at high magnetic latitudes.

As an illustration, we apply these methods to estimate the magnetization direction of a synthetic dipping slab shown in **Figure 41(a)**. The magnetization of the dipping slab is in the direction of $(I_M, D_M) = (45°, 75°)$, while the inducing field direction is $(I, D) = (65°, -25°)$. Using the cross correlation between the first vertical derivative of the RTP field and the total gradient, the estimated magnetization direction is (45°, 80°) with the correlation map shown in **Figure 42**.

The comparison of results with wavelet and the Helbig methods is shown in **Table 3**. We quantify the error by calculating the angle between the true direction and the estimated direction for each method (listed as deviation in the Table). The maximum error for this example is 9.8°. This is well within the error tolerance of 15° noted by Shearer (2005). Therefore, anyone of these estimates can be used in the subsequent inversion.

11.11.8.2 Amplitude Inversion

As discussed in **Section 11.11.4.9**, amplitude data have a weak dependence on the direction of magnetization. Therefore, one can formulate an inverse problem to recover the magnitude of the magnetization distribution in the subsurface (Li et al., 2010; Shear, 2005). This method was originally developed to tackle remanent magnetization, but studies have also shown that amplitude inversion is applicable to cases with strong self-demagnetization effect (e.g., Krahenbuhl et al., 2011; Santos and Li, 2012). Similarly to amplitude inversion, one can invert any nearly invariant quantity that can be computed from

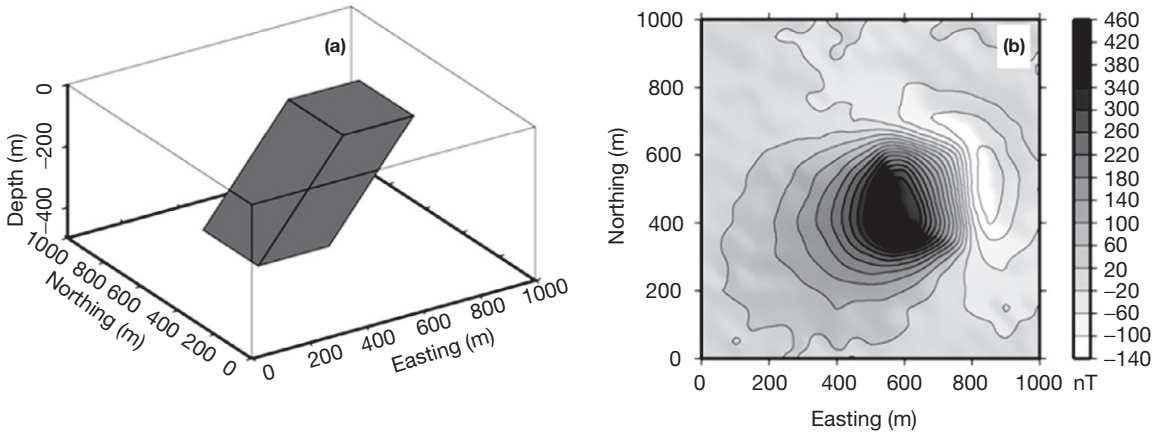


Figure 41 A synthetic example with strong remanent magnetization. (a) Dipping body viewed from the southwest direction. It has a total magnetization direction of $(I_m, D_m) = (45°, 75°)$, which deviates significantly from the inducing field direction of $(I, D) = (65°, -25°)$. (b) The total-field anomaly with a 5 nT random noise.

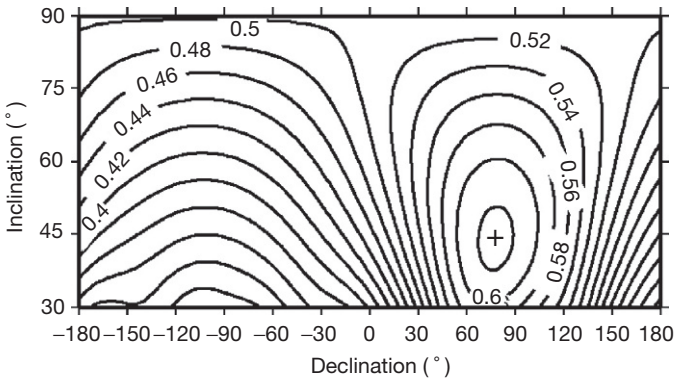


Figure 42 Cross correlation map between the total gradient of and the vertical gradient of the RTP field computed from assumed inclination and declination of the magnetization. The cross correlation achieves maximum at (45°, 80°), which is a good approximation to the true values. Reproduced from Dannemiller N and Li Y (2006) A new method for estimation of magnetization direction. *Geophysics* 71(6): L69–L73.

Table 3 Magnetization direction estimated using three different methods for the dipping slab data set shown in **Figure 41(b)**

Method	*Inclination (°)*	*Declination (°)*	*Deviation (°)*
True value	45	75	
Helbig method	46	61	9.8
Wavelet method	49	68	6.2
Cross correlation	45	80	3.5

The deviation is defined as the angle between the true and the estimated directions in 3-D.

observed magnetic data. One such quantity is the normalized source strength presented by Beiki et al. (2012). Pilkington and Beiki (2013) developed such an inversion algorithm and presented excellent illustrations. In contrast to the direction estimation method, one advantage of this class of approach is that they are not limited to a single anomaly nor do they require that the adjacent anomalies have the same magnetization direction. Therefore, the approach is generally applicable to a wide range of problems where the source distribution is more complicated. In this section, we present the basics of an amplitude inversion algorithm. The inversion of other similar nearly invariant quantities can be carried out in the same manner, as long as one can perform the forward modeling and sensitivity.

The amplitude data are defined by the following formula:

$$B_a = |\mathbf{B}_a| = \sqrt{B_{ax}^2 + B_{ay}^2 + B_{az}^2} \qquad [122]$$

where B_a is the amplitude and (B_{ax}, B_{ay}, B_{az}) are the three orthogonal components of the magnetic anomaly. The inversion algorithm starts by calculating the amplitude of the anomalous magnetic field from the observed total-field anomaly. This is accomplished by first transforming the total-field anomaly into the three orthogonal components in x-, y-, and z-directions. The amplitude data are given by eqn [122] where (B_{ax}, B_{ay}, B_{az}) are the transformed magnetic anomaly vectors. A common approach to obtain the three orthogonal components is to use the wave number-domain expressions (e.g., Pedersen, 1978) when the data are located on a plane. Alternatively, equivalent source techniques (Dampney, 1969) can be used to carry out the transformation when ground data are acquired in areas with high topographic relief, and therefore, the wave number-domain approach, which assumes that all observations lie on a planar surface, is not applicable.

To carry out the inversion, we use a similar approach to that in the magnetic susceptibility inversion (**Section 11.11.7**). The converted amplitude data are the input data to the inversion. We use the effective susceptibility defined earlier as the unknown model to be recovered. For computational purposes, we adopt a commonly used model representation that discretizes the model region in 3-D into a set of contiguous rectangular prisms by an orthogonal mesh and assume a constant effective susceptibility value within each prism.

With the representation described earlier, each component of the magnetic anomaly vector is given by a matrix–vector product in exactly the same way as in the total-field case:

$$\vec{d}_x = \mathbf{G}_x \vec{\kappa}$$

$$\vec{d}_y = \mathbf{G}_y \vec{\kappa}$$

$$\vec{d}_z = \mathbf{G}_z \vec{\kappa} \qquad [123]$$

where $\vec{d}_x = (B_{x1}, \ldots, B_{xN})^T$ is an algebraic vector holding the x-components of the anomalous magnetic field, $\vec{d}_y$ and $\vec{d}_z$ are similarly defined, $\vec{\kappa} = (\kappa_1, \ldots, \kappa_M)^T$ is the vector of unknown effective susceptibility to be recovered, and $\mathbf{G}_x$, $\mathbf{G}_y$, and $\mathbf{G}_z$ are the sensitivity matrices relating the respective components of anomalous magnetic field to effective susceptibilities. The elements of the sensitivity matrices quantify the field produced at the ith observation location by a unit effective susceptibility in the jth prism.

Combining eqns [122] and [123] yields the requisite forward mapping that relates the amplitude data to the effective susceptibility model. This relationship is nonlinear and we have a nonlinear problem regardless of other constraints.

The basic inversion algorithm follows that of the susceptibility inversion described in **Section 11.11.7**, in which the Tikhonov regularization formalism is used to trade off between the data misfit and the structural complexity of the recovered model. The data misfit, model objective function, and bound-constrained minimization problem are the same as those in eqns [101], [102], and [105], respectively, for the magnetic susceptibility inversion. One only needs to replace the

amplitude data for the total-field anomaly and the effective susceptibility for the susceptibility. The solution is obtained iteratively since nonlinearity is introduced by both the positivity constraint and the nonlinear relationship between amplitude data and effective susceptibility. A major difference, however, is that the Jacobian matrix must be computed for each iteration because of the nonlinear dependence of amplitude data on effective susceptibility.

The Newton method can be used to perform the minimization. Assuming the susceptibility model is $\boldsymbol{\kappa}^{(n)}$ at the nth iteration, we solve for a model perturbation from the following linearized system of equations:

$$\left(J^T W_d^T W_d J + \beta Z^T W_m^T W_m Z + \lambda^{(n)} X_l^{-2} + \lambda^{(n)} X_u^{-2}\right) \Delta \vec{\kappa} = J^T W_d^T W_d \left(\vec{d}^{\,\text{obs}} - \vec{d}^{\,(n)}\right) + \beta Z^T W_m^T W_m Z \left(\vec{\kappa}_0 - \vec{\kappa}\right) - \lambda^{(n)} \left(X_l^{-1} + X_u^{-1}\right) \vec{e} \qquad [124]$$

where $\vec{d}^{\,\text{obs}} = \left(B_{a1}^{\text{obs}}, \cdots, B_{aN}^{\text{obs}}\right)^T$ are the amplitude data computed from observed total-field anomaly, $\vec{d}^{\,(n)} = \left(B_{a1}^{(n)}, \cdots, B_{aN}^{(n)}\right)^T$ are the predicted amplitude data, and J is the Jacobian matrix. The remaining quantities have the same meaning as in eqn [111]. Once the model perturbation is obtained, the model update and the iteration continue exactly the same way as in magnetic susceptibility inversion.

The key is then how to obtain the Jacobian matrix, whose elements are the partial derivative of the amplitude data with respect to the effective susceptibility. We return to eqns [122] and [123]. Given that assuming the susceptibility model is $\vec{\kappa}^{(n)}$ at the nth iteration, substituting eqn [123] into eqn [122] and differentiating B_{ai} with respect to κ_j yields

$$\frac{\partial B_{ai}}{\partial \kappa_j} = \frac{\vec{B}_{ai}^{\,(n)}}{\left|\vec{B}_{ai}^{\,(n)}\right|} \cdot \vec{b}_{ij} \qquad [125]$$

where $\vec{B}_{ai}^{\,(n)}$ is the predicted anomalous magnetic vector at the ith observation location by the model $\vec{\kappa}^{(n)}$ at the nth iteration and $\vec{b}_{ij}$ is the magnetic vector produced at the same location by a unit susceptibility in the jth prism. The three components of $\vec{b}_{ij}$ are, respectively, the elements of the three sensitivity matrices in eqn [123].

Thus, the Jacobian has a simple form given by the inner product of the unit vector of the predicted magnetic field at the nth iteration and the magnetic field produced by a unit effective susceptibility in a prism. The three components of the latter are, respectively, the elements of the three sensitivity matrices in eqn [123]. For computational purpose, this means that we only need to compute and store the three sensitivity matrices and calculate the sensitivity for the amplitude data by a sequence of matrix–vector multiplications when solving eqn [124].

When calculating the three individual sensitivity matrices, we must formally specify a magnetization direction. Given that we are utilizing the weak dependence of amplitude data on magnetization direction, it is sufficient to use the direction of the current inducing field. Thus, the sensitivity is computed as if the magnetization vector is given by the product of the effective susceptibility and the inducing field. It is important to note, however, that the interpretation of the inversion result is independent of the direction of the inducing field.

Because the iterative inversion relies on the three linear sensitivity matrices, and we impose same type of bound constraints, we can transfer all the computational techniques developed in the magnetic susceptibility to amplitude inversion. These include the use of primal logarithmic barrier method, the wavelet compression of sensitivity matrices, and the use of CG method to solve the linear system equations. In terms of total computational cost, the amplitude inversion requires approximately three times more time than the inversion of total-field data for susceptibility.

11.11.8.2.1 Synthetic example

As illustration, we invert the total-field data shown in Figure 41 (b) first by using the estimated magnetization direction and then by converting them into amplitude data and applying the amplitude inversion.

For the former, we use the estimation from the wavelet method ($I_m = 49^\circ$ and $D_m = 68^\circ$), whose deviation from the true direction is the middle of the three estimation results from Table 3. We use the algorithm described in the Section 11.11.7 and explicitly specify the magnetization direction as that estimated using the wavelet method (Table 3). The result (Figure 43, left column) clearly shows the dip of the source and its horizontal and vertical extent. Over all, the recovered anomalous body is a good representation of the true model.

The result of amplitude inversion is shown in the right column of Figure 43, which can be compared with the model recovered using an estimated magnetization direction (left column). We note that the location and the spatial extent of the anomalous body are recovered to some extent, but the dip is less visible. Consequently, the source body center is located deeper. This result is to be expected. Since much of the dip information is encoded in the phase of the magnetic anomaly and that information is lost in the amplitude data. Despite the lack of the dip information, the inversion result is a much improved one compared to the inability to invert the original data without the magnetization direction when using magnetic susceptibility inversion algorithms.

11.11.8.3 Magnetization Vector Inversion

The primary difficulty with the inversion of magnetic data affected by remanent magnetization is the unknown magnetization direction. The approaches discussed in the two preceding subsections seek either to estimate the direction prior to inversion or to formulate the inverse problem to circumvent the need for the direction. Both approaches produce the magnitude of the magnetization vector as a function of the spatial position in the end. The outcome would be to recover the magnetization direction with the magnitude simultaneously. This is an active area of research and several algorithms have been developed.

Wang et al. (2004) presented such an algorithm suitable for determining the total magnetization direction of separated, homogeneous bodies. Lelièvre and Oldenburg (2009) presented

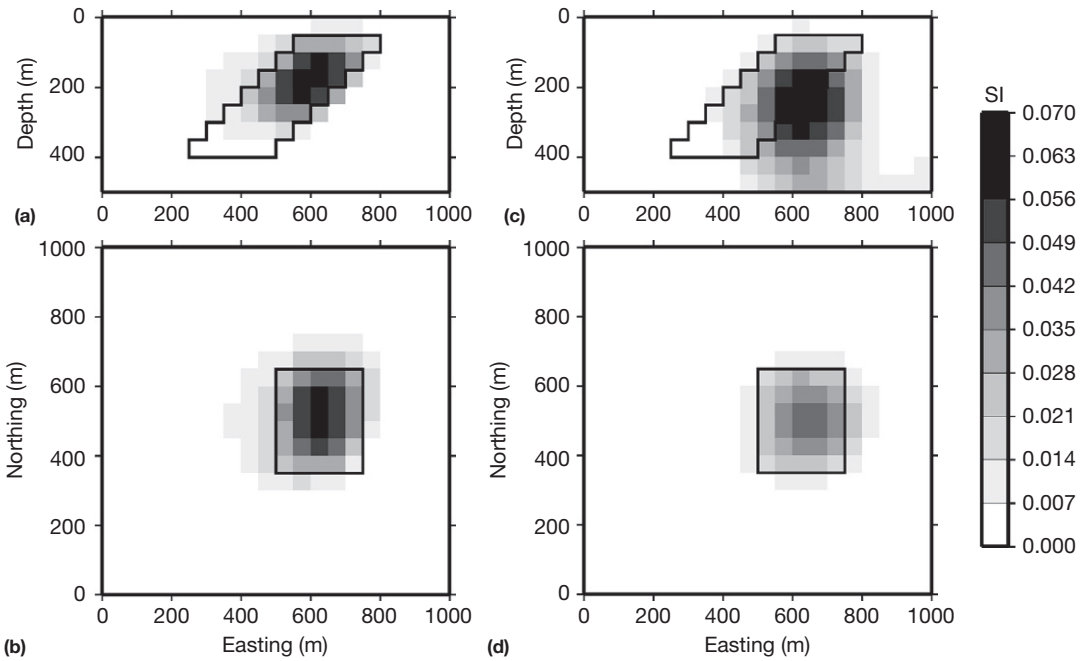


Figure 43 Comparison of inversion result obtained by using the estimated magnetization direction (a and b) and that by inverting amplitude data directly (c and d).

a comprehensive study of magnetization vector inversion (MVI), in which they discussed two alternative parameterizations and explored the need to impose geologic constraints to produced interpretable magnetization vector models. Ellis et al. (2012) presented a similar algorithm and employed a compactness model objective function on magnitude of magnetization. Liu et al. (2013) developed an MVI for borehole data to image highly magnetic ore bodies located between multiple drill holes.

We present the basics of this method here. We assume a set of total-field magnetic anomaly data $\vec{d}^{\,\text{obs}} = \left(\Delta T_1^{\text{obs}}, \ldots, \Delta T_i^{\text{obs}}, \ldots\right)^T$ produced by a 3-D distribution of magnetization vector $\vec{M}(x, y, z)$. We represent the magnetization by a set of contiguous rectangular cells as in the magnetic susceptibility or amplitude inversions and assume that each cell has a constant magnetization vector $\vec{M}_j = \left(M_{xj}, M_{yj}, M_{zj}\right)$. The inverse problem of MVI seeks to reconstruct the 3-D distribution of $\vec{M}(x, y, z)$ from the total-field anomaly, possibly with any constraints formed from prior information.

For convenience, we define the model for the inverse problem as the algebraic vector formed by concatenating the three components of the magnetization vectors:

$$\vec{m} = \begin{pmatrix} \vec{M}_x \\ \vec{M}_y \\ \vec{M}_z \end{pmatrix} \qquad [126]$$

where

$$\begin{aligned} \vec{M}_x &= \left(M_{x1}, \ldots, M_{xj}, \ldots\right)^T \\ \vec{M}_y &= \left(M_{y1}, \ldots, M_{yj}, \ldots\right)^T \\ \vec{M}_z &= \left(M_{z1}, \ldots, M_{zj}, \ldots\right)^T \end{aligned} \qquad [127]$$

are three model vectors corresponding to three magnetization components. The forward modeling is then given by

$$\left[G_x, G_y, G_z\right]\vec{m} = \vec{d} \qquad [128]$$

where G_x, G_y, and G_z are, respectively, the sensitivity matrices of the total-field anomaly with respect to the x-, y-, and z-components of the magnetization vector. These three matrices are similar to those used in the amplitude inversion in the preceding subsection, except for a scaling factor and the specification of source direction and anomaly projection direction. The difference in scaling factor arises because the amplitude inversion was formulated using effective susceptibility as model, whereas eqn [128] uses magnetization as model.

The basic formulation of MVI uses the Tikhonov regularization and minimizes the following total objective function:

$$\varphi\left(\vec{m}\right) = \varphi_d\left(\vec{m}\right) + \beta\varphi_m\left(\vec{m}\right) \qquad [129]$$

where the data misfit $\varphi_d\left(\vec{m}\right)$ is defined as in preceding sections and $\varphi_m\left(\vec{m}\right)$ is a model objective function designed for the vector model. The two common elements it has with magnetic susceptibility inversion are the measure of model complexity on each component and the depth-weighting function. This is the essence of the algorithm presented by Lelièvre and Oldenburg (2009). However, they chose to represent the magnetization in the reference frame aligned with the inducing field direction with either three orthogonal Cartesian components or spherical representation based on magnitude and directions.

We note that the MVI attempts to recover three 3-D functions (e.g., the three components of the magnetization) from a 2-D magnetic data set. Therefore, the associated ambiguity is much more severe than that encountered in the inversion for

magnetic susceptibility or magnitude of magnetization, both of which deal with one 3-D function. Thus, additional constraints are required for MVI. Variations of such constraints and their implementation then give rise to different MVI algorithms.

Lelièvre and Oldenburg (2009) discussed in detail how to impose various prior data such as the known total magnetization direction, measured magnetic susceptibility, measured Königsberger ratio (Q), or estimated remanent magnetization direction. Ellis et al. (2012) imposed a compactness objective function related to an L_0 norm of the magnitude of magnetization. Li and Sun (2013) imposed constraints on the range of variation of magnetization direction by the fuzzy c-means clustering technique. Much work remains to be done, and we expect to see both existing algorithms to mature and new algorithms to emerge.

11.11.8.3.1 A synthetic example

An example of MVI is shown in Figure 44 (Lelièvre and Oldenburg, 2009). The model consists of two parallel rectangular source bodies with different magnetization directions in NW and NE direction, respectively. Because of the proximity of the two sources, the total-field anomaly appears to merge together (Figure 44(a)). The total MVI, with constraints that limit the extent of the source region, is able to recover the compact source region of high magnetization strength (Figure 44(b)) and identify the different magnetization directions dominant in the west and east portions of the source. This example also illustrates the importance and potential utility of magnetization direction recovered from the inversion. Although the recovered magnetization strength does not indicate the presence of two separate causative bodies in this example, the magnetization directions clearly show that two different source regions are present. The authors' application of their algorithm to a field data set also demonstrates the similar advantage.

11.11.8.4 A Field Example

As a field example, we present an example data set from diamond exploration (Figure 44(a)), which was acquired by Teck Cominco and Diamonds North over kimberlites on Victoria Island in the Northwest Territories, Canada (Li et al., 2010). The geology in the area is characterized by Archean granitic basement overlain by a Proterozoic sedimentary sequence with minor volcanics, capped by flat-lying Cambrian-to-Devonian carbonate rocks. The host rocks are largely nonmagnetic, and the kimberlite bodies stand out in the magnetic surveys as distinct 'sharp' anomalies indicative of shallow bodies, compared with the more rounded anomalies caused by deep features in the basement or within the sedimentary rocks. Both positive and negative anomalies are associated with kimberlite intrusions, whereas the negative anomalies are produce by hypabyssal dikes. The ages of the kimberlite intrusions are in the 250–300 million-year range. The distinct magnetic signature of these kimberlite dikes above a quiet magnetic background makes high-resolution magnetic surveys an ideal exploration tool in this area. However, the highly variable orientation of these anomalies is produced by a magnetization that is dominated by strong remanent magnetization that has variable directions. As a result, the quantitative interpretation of these anomalies is difficult to achieve through magnetic susceptibility inversion algorithms.

The data show a number of dipolar total-field anomalies with differing orientations throughout the survey area. The map has been rotated clockwise by 34° and the inducing field has an inclination of 86.7° and a nominal declination of −7.7°. Two types of anomalies dominate the data set: several small, compact, and high-frequency anomalies surround areas of broad, lower-frequency anomalies. It is clear from the orientation of the anomalies that the total magnetization direction varies greatly from anomaly to anomaly. Thus, it is unlikely to be able to invert this data set with a single magnetization direction. We resort to the second approach, that is, we

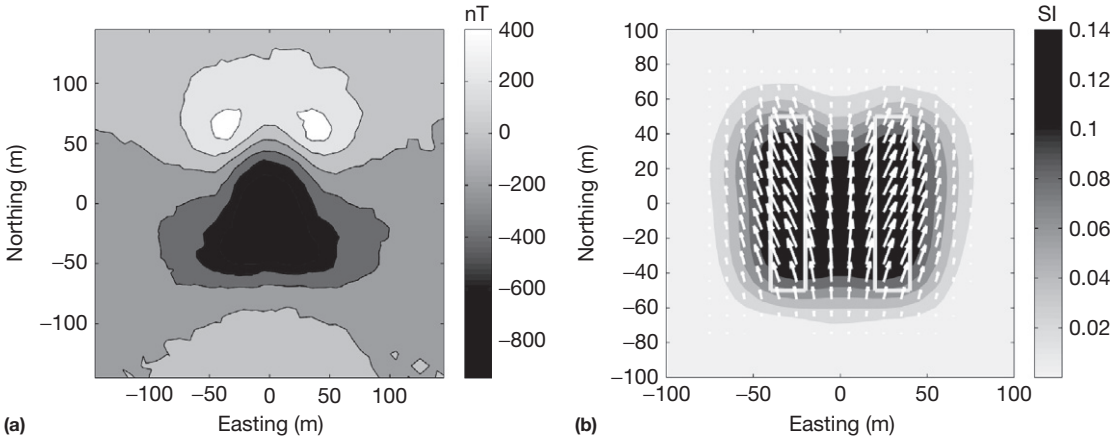


Figure 44 A synthetic example with two source bodies having differing magnetization directions. (a) The total-field anomaly. (b) A plan section at 30 m depth through the recovered 3-D magnetization vector model. The color contour shows the recovered effective susceptibility, white arrows show the inverted magnetization direction, and the lengths of arrows are proportional to the effective susceptibility. The gray rectangles outline the position of the source bodies. Reproduced from Lelièvre PG and Oldenburg DW (2009) A 3D total magnetization inversion applicable when significant, complicated remanence is present. *Geophysics* 74(3): L21–L30.

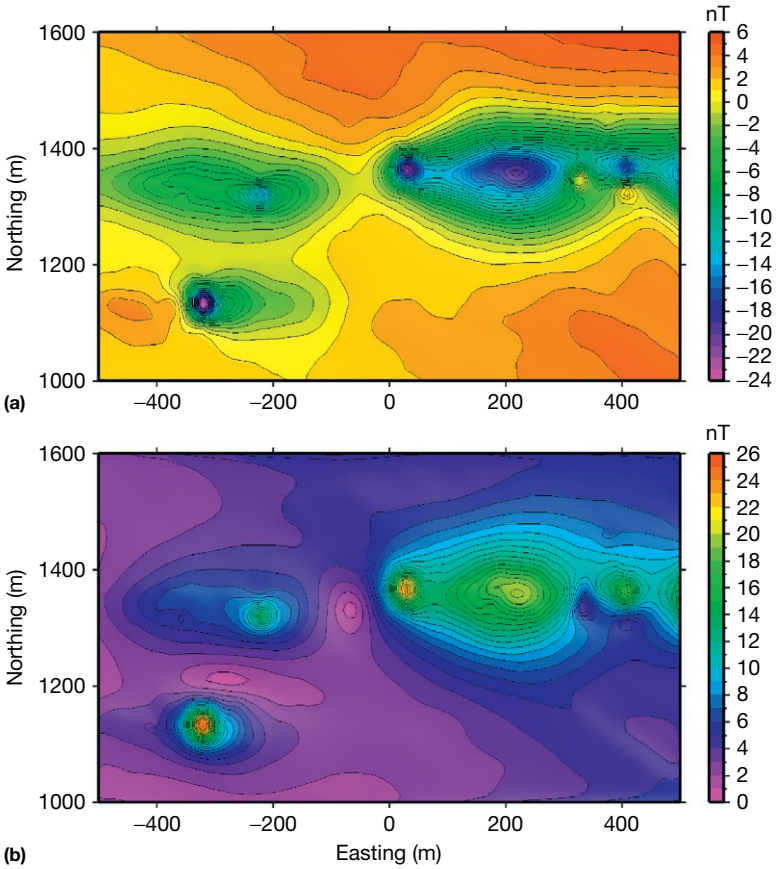


Figure 45 (a) The total-field anomaly data over a group of kimberlites on Victoria Island, NWT, Canada. The 600 m by 1000 m data have been rotated 34° clockwise and regridded at a 10 m spacing. The rotated data map has an inducing field with an inclination of 86.7° and a nominal declination of −7.7°. (b) The corresponding amplitude of anomalous magnetic vector.

choose to invert the amplitude of the anomalous magnetic field and recover the magnitude of magnetization in the form of an effective susceptibility. The computed amplitude data are shown in Figure 45(b).

The effective susceptibility recovered from the inversion of the amplitude data in Figure 45(b) is shown in Figure 46 as a volume-rendered image. The isosurfaces correspond to the effective susceptibility value of 0.0012 SI. There are five main anomalous bodies of high susceptibility in the recovered model, correlating well with the five anomalies in the total-field and amplitude data. The broad elongated magnetic body resembles a kimberlite dike known in this area, whereas the more compact bodies oriented vertically correspond to kimberlite pipes.

11.11.9 Discussion

Magnetic data are one of the most widely acquired and used types of geophysical data. They have played an important role in a broad range of problems such as mineral and energy exploration, crustal studies, and environmental and archaeological investigations. It is not surprising that a large number of methods and algorithms have been developed to process magnetic data and to extract from them information about subsurface targets and structures. These methods have evolved to cover a wide spectrum ranging from enhancement of magnetic data images for target and structure identification, to various source parameters extraction methods, to quantitative interpretation based on inversions for geologic boundaries and for 3-D distribution of susceptibility or magnetization in the subsurface. Given the variety of problems that employ magnetic data and the varying state of practice, nearly all these methods are used to some extent.

For regional- and reconnaissance-scale problems, methods such as depth estimation and boundary extraction have been ubiquitous. These methods will continue to play an important role for such large-scale data sets and for a first-order interpretation. Meanwhile, for applications such as imaging and characterizing specific structures or targets, inversions are becoming the methods of choice. With the improvement of existing algorithms and the development of new ones, as well as the rapidly increasing computing power, the use of magnetic inversion will increase and its areas of application will expand. Within the limited scope of this chapter, we have endeavored to present the magnetic methods in this context: focusing on those that we feel are most commonly used or whose usage will expand significantly as we move forward.

Looking back at the historical development of magnetic methods, it has been one of steady advancement from

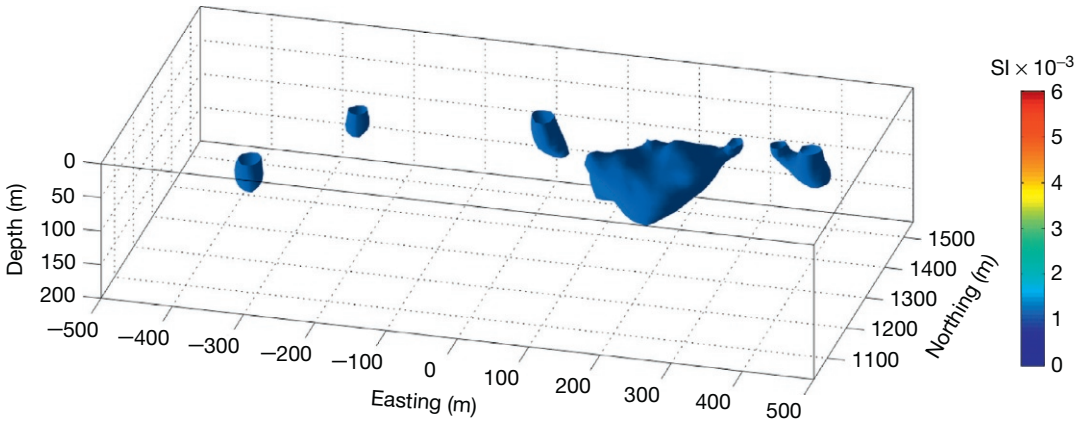


Figure 46 Volume-rendered recovered effective susceptibility obtained from inverting the amplitude data shown in **Figure 45(b)**. The view is from the southeast. The color scale indicates the full range of effective susceptibility. The display shows isosurfaces at 0.0012 SI value. Five magnetic sources are recovered.

qualitative to quantitative approaches. If the past could be described as 'anomaly bump hunting,' the present would be 3-D imaging of the subsurface in which we strive to define the distribution of magnetic properties in various geologic units. With the availability of 3-D magnetic property models from inversion, the interpretation has moved from the data domain to the model domain. Albeit a simplistic characterization, the present could be described as 'magnetic property bump hunting.' Looking ahead, it may be argued that the next stage of development is the quantitative use of recovered magnetic models for differentiating and characterizing geology and geologic processes. Such efforts are already being pursued.

In order to accomplish geology differentiation and characterization, magnetic data must be integrated with other geophysical data and with petrophysical and geologic data. Therefore, we anticipate significant research and application of joint inversions, and integrated interpretation, of magnetic data with other geoscientific data sets in the near future.

Acknowledgments

We thank Lee Slater for inviting us to write this chapter and for his effort in editing and handling the manuscript. We thank the members of the Center for Gravity, Electrical & Magnetic Studies (CGEM) for the many discussions and help. Leon Foks, Cericia Martinez, and Hyoungrea Rim helped refine many figures in the chapter. Joe Capriotti, Camriel Coleman, Leon Foks, Susana Guzman, Cericia Martinez, and Jiajia Sun provided many helpful comments on the manuscript. We also thank Mark Pilkington for his comments and suggestions. We finally thank Kris Davis and Peter Lelièvre for providing some of the figures.

References

Aharoni A (2000) Introduction to the theory of ferromagnetism. *International Series of Monographs on Physics*, 2nd edn., 93 Oxford: Oxford Science Publications.

Arkani-Hamed J (1988) Differential reduction to the pole of regional magnetic anomalies. *Geophysics* 53: 1592–1600.

Ballantyne EJ (1980) Magnetic curve fit for a thin dike: Calculator program (TI 59). *Geophysics* 45: 447–455.

Baranov V (1957) A new method for interpretation of aeromagnetic maps pseudo-gravimetric anomalies. *Geophysics* 22: 359–383.

Baranov V and Naudy H (1964) Numerical calculation of the formula of reduction to the magnetic pole. *Geophysics* 29: 67–79.

Barbosa VCF and Silva JBC (2011) Reconstruction of geologic bodies in depth associated with a sedimentary basin using gravity and magnetic data. *Geophysical Prospecting* 59: 1021–1034.

Barnett CT (1976) Theoretical modeling of the magnetic and gravitational fields of an arbitrarily shaped three-dimensional body. *Geophysics* 41: 1353–1364.

Barrett R, Berry MW, Chan TF, et al. (1993) *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. Philadelphia, PA: SIAM.

Beiki M, Clark DA, Austin JR, and Foss C (2012) Estimating source location using normalized magnetic source strength calculated from magnetic gradient tensor data. *Geophysics* 77(6): J23–J37.

Bhattacharyya BK (1964) Magnetic anomalies due to prism-shaped bodies with arbitrary polarization. *Geophysics* 29: 517–531.

Bhattacharyya BK (1965) Two-dimensional harmonic analysis as a tool for magnetic interpretation. *Geophysics* 30: 829–857.

Bhattacharyya BK (1966) Continuous spectrum of the total-magnetic-field anomaly due to a rectangular prismatic body. *Geophysics* 31: 97–121.

Bhattacharyya BK (1980) A generalized multibody model for inversion of magnetic anomalies. *Geophysics* 45: 255–270.

Blakely RJ (1995) *Potential Theory in Gravity and Magnetic Applications*. Cambridge: Cambridge University Press.

Blakely RJ (1996) *Potential Theory in Gravity and Magnetic Applications*. Cambridge: Cambridge University Press, 441 pp.

Blakely RJ and Simpson RW (1986) Approximating edges of source bodies from magnetic or gravity anomalies. *Geophysics* 51(7): 1494–1498.

Born M (1999) *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th edn. Cambridge University Press.

Bott MHP (1963) Two methods applicable to computers for evaluating magnetic anomalies due to finite three dimensional bodies. *Geophysical Prospecting* 11: 292–299.

Burger HR (1992) *Exploration Geophysics of the Shallow Subsurface*: Prentice Hall.

Byerly PE (1965) Convolution filtering of gravity and magnetic maps. *Geophysics* 30: 281–283.

Cady JW (1980) Calculation of gravity and magnetic anomalies of finite-length right polygonal prisms. *Geophysics* 45: 1507–1512.

Chapin DA (1997) Wavelet transforms: A new paradigm for interpreting gravity and magnetics data. In: *67th Annual International Meeting, SEG, Expanded Abstracts*, pp. 486–489.

Cooper GRJ and Cowan DR (2008) Edge enhancement of potential-field data using normalized statistics. *Geophysics* 73(3): H1–H4.

Cordell L (1979) Gravimetric expression of graben faulting in Santa Fe County and the Espanola Basin, New Mexico. In: Ingersoll RV (ed.). *Guidebook to Santa Fe County: 30th Field Conference, New Mexico Geological Society*, pp. 59–64.
Cordell L (1992) A scattered equivalent-source method for interpolation and gridding of potential-field data in three dimensions. *Geophysics* 57(4): 629–636.
Cordell L and McCafferty AE (1989) A terracing operator for physical property mapping with potential field data. *Geophysics* 54: 621–634.
Cowan DR and Cowan S (1993) Separation filtering applied to aeromagnetic data. *Exploration Geophysics* 24: 429–436.
Cox A (1969) Geomagnetic reversals. *Science* 163: 237–245.
Cribb J (1976) Application of the generalized linear inverse to the inversion of static potential field data. *Geophysics* 41: 1365–1369.
Dampney CNG (1969) The equivalent source technique. *Geophysics* 34(1): 39–53.
Dannemiller N and Li Y (2006) A new method for estimation of magnetization direction. *Geophysics* 71(6): L69–L73.
Daubechies I (1992) *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM.
Davis K, Haber E, and Oldenburg DW (2013) Large-scale magnetic inversion using differential equations and ocTrees. In: *23rd Geophysical Conference and Exhibition, Australian Society of Exploration Geophysics, Melbourne, VIC, 11–14 August*.
Davis K and Li Y (2011) Fast solution of geophysical inversion using adaptive mesh, space-filling curves, and wavelet compression. *Geophysical Journal International* 185: 157–166.
Davis K and Li Y (2012) Joint processing of total-field and gradient magnetic data using equivalent source technique. *Exploration Geophysics* 42: 199–206.
Davis K and Li Y (2013) Efficient 3D inversion of magnetic data via octree-mesh discretization, space-filling curves, and wavelets. *Geophysics* 78(5): J61–J73.
Davis K, Li Y, and Nabighian M (2010) Automatic detection of UXO magnetic anomalies using extended Euler deconvolution. *Geophysics* 75(3): G13–G20.
Dean WC (1958) Frequency analysis for gravity and magnetic interpretation. *Geophysics* 23: 97–127.
Dey A and Morrison HF (1979) Resistivity modelling for arbitrarily shaped three-dimensional structures. *Geophysics* 44: 753–780.
Dransfield M, Christensen A, and Liu G (2003) Airborne vector magnetic mapping of remanently magnetized banded iron formations at Rocklea, Western Australia. *Exploration Geophysics* 34: 93–96.
Ellis R, de Wet B, and Macleod IM (2012) Inversion of magnetic data from remanent and induced sources. In: *Presented at the 22nd ASEG International Geophysical Conference*.
Emilia DA (1973) Equivalent sources used as an analytic base for processing total magnetic field profiles. *Geophysics* 38: 339–348.
Eskola T, Tervo T, and Puranen R (1977) Calculation of magnetostatic anomalies and demagnetization factor by means of the method of subareas. *Geological Survey of Finland*. Report of Investigation 16.
Foss C and McKenzie B (2011) Inversion of anomalies due to remanent magnetisation: An example from the Black Hill Norite of South Australia. *Australian Journal of Earth Sciences* 58: 391–405.
Fuller BD (1967) Two-dimensional frequency analysis and design of grid operators. In: *Mining Geophysics*, pp. 658–708. Tulsa, OK: Society of Exploration Geophysicists, II.
Gerovska D, Arauzo-Bravo MJ, and Stavrev P (2009) Estimating the magnetization direction of sources from southeast Bulgaria through correlation between reduced-to-the-pole and total magnitude anomalies. *Geophysical Prospecting* 57: 491–505.
Gill PE, Murray W, Ponceleon DB, and Saunders M (1991) Solving reduced KKT systems in barrier methods for linear and quadratic programming. Technical Report SOL 91–7. Stanford, CA: Stanford University.
Goupillaud PL, Grossmann A, and Morlet J (1984) Cycle-octave and related transforms in seismic signal analysis. *Geoexploration* 23: 85–102.
Grant FS and West GF (1965) *Interpretation Theory in Applied Geophysics*. New York: McGraw-Hill.
Guillen A and Menichetti V (1984) Gravity and magnetic inversion with minimization of a specific functional. *Geophysics* 49: 1354–1360.
Gunn PJ (1972) Application of Wiener filters to transformations of gravity and magnetic fields. *Geophysical Prospecting* 20: 860–871.
Gunn PJ (1995) An algorithm for reduction to the pole that works at all magnetic latitudes. *Exploration Geophysics* 26: 247–254.
Haney M, Johnston C, Li Y, and Nabighian MN (2003) Envelopes of 2D and 3D magnetic data and their relationship to the analytic signal: Preliminary results. In: *73rd Annual International Meeting, SEG Expanded Abstracts*, pp. 596–599.
Haney M and Li Y (2002) Total magnetization direction and dip from multiscale edges. In: *72nd Annual International Meeting, SEG, Expanded Abstracts*, pp. 735–738.
Hansen PC (1992) Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review* 34: 561–580.
Hansen RO and Miyazaki Y (1984) Continuation of potential fields between arbitrary surfaces. *Geophysics* 49(6): 787–795.
Hansen RO and Pawlowski RS (1989) Reduction-to-the- pole at low latitudes by Wiener filtering. *Geophysics* 54: 1607–1613.
Hansen RO and Simmonds M (1993) Multiple-source Werner deconvolution. *Geophysics* 58: 1792–1800.
Hansen RO and Suciu L (2002) Multiple-source Euler deconvolution. *Geophysics* 67: 525–535.
Hardwick CD (1999) Gradient-enhanced total-field gridding. *69th Annual International Meeting Society of Exploration Geophysics, Expanded Abstracts*, 18, pp. 381–385.
Harrington RF (1968) *Field Computation by Moment Methods*. New York: The Macmillan Company.
Hartman RR, Teskey DJ, and Friedberg JL (1971) A system for rapid digital aeromagnetic interpretation. *Geophysics* 36: 891–918.
Heirtzler JR, LePichon X, and Baron JG (1966) Magnetic anomalies over the Reykjanes Ridge. *Deep Sea Research* 13: 427–432.
Helbig K (1962) Some integrals of magnetic anomalies and their relation to the parameters of the disturbing body. *Zeitschrift für Geophysik* 29: 83–96.
Helbig K (1963) Some integrals of magnetic anomalies and their relation to the parameters of the disturbing body. *Zeitschrift für Geophysik* 29(2): 83–95.
Henderson RG (1960) A comprehensive system of automatic computation in magnetic and gravity interpretation. *Geophysics* 25: 560–585.
Henderson RG (1970) On the validity of the use of the upward continuation integral for total magnetic intensity data. *Geophysics* 35(5): 916–919.
Hestenes MR and Stiefel E (1952) Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards* 47: 409–436.
Hestenes MR and Stiefel E (1970) On the validity of the use of the upward continuation integral for total magnetic intensity data. *Geophysics* 35: 916–919.
Hilbert D (1891) Ueber die stetige abbildung einer line auf ein flächenstück. *Mathematische Annalen* 38(1): 459–460.
Hjelt SE (1972) Magnetostatic anomalies of a dipping prism. *Geoexploration* 10: 239–254.
Holstein H (2002a) Gravimagnetic similarity in anomaly formulas for uniform polyhedra. *Geophysics* 67: 1126–1133.
Holstein H (2002b) Invariance in gravimagnetic anomaly formulas for uniform polyhedra. *Geophysics* 67: 1134–1137.
Hood P (1965) Gradient measurements in aeromagnetic surveying. *Geophysics* 30: 891–902.
Hornby P, Boschetti F, and Horowitz FG (1999) Analysis of potential field data in the wavelet domain. *Geophysical Journal International* 137: 175–196.
Jacobsen B (1987) A case for upward continuation as a standard separation filter for potential-field maps. *Geophysics* 52: 1138–1148.
Jessell MW (2001) Three-dimensional modeling of potential-field data. *Computers and Geosciences* 27: 455–465.
Jessell MW and Valenta RK (1996) Structural geophysics: Integrated structural and geophysical mapping. In: DePaor D (ed.) *Structural Geology and Personal Computers*, pp. 303–324. Elsevier Science Publishing Co.
Keating P and Zerbo L (1996) An improved technique for reduction to the pole at low latitudes. *Geophysics* 61: 131–137.
Kellogg OD (1953) *Foundations of Potential Theory*. Dover Publications.
Krahenbuhl R and Li Y (2007) Influence of self-demagnetization effect on data interpretation in strongly magnetic environments. *SEG Technical Program Expanded Abstracts 2007*. pp. 713–717.
Krahenbuhl R, Li Y, Nabighian M, Davis K, and Billings S (2011) Advanced UXO detection and discrimination using magnetic data based on extended Euler deconvolution and shape identification through multipole moments. *Project MR-1638 Final Report*, *Strategic Environmental Research and Development* (*SERDP*).
Lee TJ (1980) Rapid computation of magnetic anomalies with demagnetization included, for arbitrarily shaped magnetic bodies. *Geophysical Journal of the Royal Astronomical Society* 60: 67–75.
Lelièvre PG (2005) *Forward Modelling and Inversion of Geophysical Magnetic Data*. MSc Thesis, University of British Columbia.
Lelievre PG and Oldenburg DW (2006) Magnetic forward modeling and inversion for high susceptibility. *Geophysical Journal International* 166: 76–90.
Lelièvre PG and Oldenburg DW (2009) A 3D total magnetization inversion applicable when significant, complicated remanence is present. *Geophysics* 74(3): L21–L30.
Li Y (2001) Processing gravity gradiometer data using an equivalent source technique. *SEG Technical Program Expanded Abstracts 2001*, pp. 1466–1469.
Li X (2006) Understanding 3D analytic signal amplitude. *Geophysics* 71(2): L13–L16.
Li Y, Devriese SGR, Krahenbuhl RA, and Davis K (2013) Enhancement of magnetic data by stable downward continuation for UXO application. *IEEE Transactions on Geoscience and Remote Sensing* 51: 3605–3614.

Li S and Li Y (2014) Inversion of magnetic anomaly on rugged observation surface in the presence of strong remanent magnetization. *Geophysics* 79: J11–J19.

Li Y and Oldenburg DW (1996) 3-D inversion of magnetic data. *Geophysics* 61: 394–408.

Li Y and Oldenburg DW (2000a) Joint inversion of surface and three-component borehole magnetic data. *Geophysics* 65: 540–552.

Li Y and Oldenburg DW (2000b) Reduction to the pole using equivalent sources. In: *60th Annual International Meeting, SEG, Expanded Abstracts*, pp. 386–389.

Li Y and Oldenburg DW (2001) Stable reduction to the pole at the magnetic equator. *Geophysics* 66: 571–578.

Li Y and Oldenburg DW (2003) Fast inversion of large-scale magnetic data using wavelet transforms and logarithmic barrier method. *Geophysical Journal International* 152: 251–265.

Li Y and Oldenburg DW (2010) Rapid construction of equivalent sources using wavelets. *Geophysics* 75(3): L51–L59.

Li Y, Shearer S, Haney M, and Dannemiller N (2010) Comprehensive approaches to 3D inversion of magnetic data affected by remanent magnetization. *Geophysics* 75(1): L1–L11.

Li Y and Sun J (2013) Total-magnetization vector inversion using statistical constraints: Guided fuzzy c-means clustering. In: *13th SAGA Biannual Conference & Exhibition*.

Liu S, Hu X, Liu T, Feng J, Gao W, and Qiu L (2013) Magnetization vector imaging for borehole magnetic data based on magnitude magnetic anomaly. *Geophysics* 78: D429–D444.

Lourenco JS and Morrison HF (1973) Vector magnetic anomalies derived from measurements of a single component of the field. *Geophysics* 38: 359–368.

Lyrio JCS, Tenorio L, and Li Y (2003) Efficient automatic denoising of gravity gradiometry data using wavelet transforms. *Geophysics* 69(3): 772–782.

Mallat SG (1989) A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7): 674–693.

Martelet G, Sailhac P, Moreau F, and Diament M (2001) Characterization of geological boundaries using 1-D wavelet transform on gravity data: Theory and application to the Himalayas. *Geophysics* 66(4): 1116–1129.

Martins CM, Lima WA, Barbosa VC, João B, and Silva JB (2011) Total variation regularization for depth-to-basement estimate: Part 1 – Mathematical details and applications. *Geophysics* 76(1): I1–I12.

Mendonca CA and Silva JBC (1993) A stable truncated series approximation of the reduction-to-the-pole operator. *Geophysics* 58: 1084–1090.

Mendonça C and Silva J (1994) The equivalent data concept applied to the interpolation of potential field data. *Geophysics* 59(5): 722–732.

Meyer Y (1993) *Wavelets—Algorithms and Applications*. Society for Industrial and Applied Mathematics.

Miller HG and Singh V (1994) Potential field tilt; a new concept for location of potential field sources. *Journal of Applied Geophysics* 32: 213–217.

Moon B, Jagadish HV, Faloutsos C, and Saltz JH (2001) Analysis of the clustering properties of Hilbert space-filling curve. *IEEE Transactions on Knowledge and Data Engineering* 13(1): 124–141.

Moreau F, Gibert D, Holschneider M, and Saracco G (1997) Wavelet analysis of potential fields. *Inverse Problems* 13: 165–178.

Mueller EL, Morris WA, Killeen PG, and Balch S (1997) Combined 3-d interpretation of airborne, surface, and borehole vector magnetics at the Mcconnell nickel deposit. In: *Presented at the 67th Annual International Meeting, Society of Exploration Geophysicists*.

Mushayandebvu MF, van Driel P, Reid AB, and Fairhead JD (2001) Magnetic source parameters of two-dimensional structures using extended Euler deconvolution. *Geophysics* 66: 814–823.

Nabighian MN (1972) The analytic signal of two-dimensional magnetic bodies with polygonal cross-section: Its properties and use for automated anomaly interpretation. *Geophysics* 37: 507–517.

Nabighian MN (1974) Additional comments on the analytic signal of two- dimensional magnetic bodies with polygonal cross-section. *Geophysics* 39: 85–92.

Nabighian MN (1984) Towards a three-dimensional automatic interpretation of potential field data via generalized Hilbert transforms: Fundamental relations. *Geophysics* 49: 780–786.

Nabighian MN (2006) A modified Helbig method, Presented at Gravity and Magnetics Research Consortium (GMRC) sponsors meeting, Golden, CO.

Nabighian MN and Hansen RO (2001) Unification of Euler and Werner deconvolution in three dimensions via the generalized Hilbert transform. *Geophysics* 66: 1805–1810.

Nakatsuka T and Okuma S (2006) Reduction of magnetic anomaly observations from helicopter surveys at varying elevations. *Exploration Geophysics* 37(1): 121–128.

Nelson BJ (1994) Levelling total-field aeromagnetic data with measured horizontal gradients. *Geophysics* 59: 1166–1170.

Nettleton LL (1942) Gravity and magnetic calculations. *Geophysics* 7: 293–310.

Nocedal J and Wright SJ (1999) *Numerical Optimization*. New York, NY: Springer Science.

Okabe M (1979) Analytical expressions for gravity anomalies due to homogeneous polyhedral bodies and translations into magnetic anomalies. *Geophysics* 44: 730–741.

Oldenburg DW, Li Y, Farquharson CG, et al. (1998) Applications of geophysical inversions in mineral exploration problems. *The Leading Edge* 17: 461–464.

Oliveira V Jr., Barbosa V, and Uieda L (2013) Polynomial equivalent layer. *Geophysics* 78(1): G1–G13.

Ontario Geological Survey (2004) Ontario airborne geophysical surveys, magnetic data, profile data (ASCII format), Central Abitibi Destor-Porcupine–Pipestone Faults area: *Ontario Geological Survey*, Geophysical Data Set 1049c.

Parker RL (1972) The rapid calculation of potential anomalies. *Geophysical Journal of the Royal Astronomical Society* 31: 447–455.

Parker RL (1977) Understanding inverse theory. *Annual Review of Earth and Planetary Sciences* 5: 35–64.

Parker RL (1994) *Geophysical Inverse Theory*. Princeton University Press.

Parker RL and Huestis SP (1974) The inversion of magnetic anomalies in the presence of topography. *Journal of Geophysical Research* 79: 1587–1593.

Pearson WC and Skinner CM (1982) Reduction-to-the-pole of low latitude magnetic anomalies. In: *52nd Annual International Meeting, SEG, Expanded Abstracts*, p. 356.

Pedersen LB (1978) Wavenumber domain expressions for potential fields from arbitrary 2, 21/2 and 3-dimensional bodies. *Geophysics* 43: 626–630.

Pedersen LB (1979) Constrained inversion of potential field data. *Geophysical Prospecting* 27: 726–748.

Peters LJ (1949) The direct approach to magnetic interpretation and its practical application. *Geophysics* 14: 290–320.

Phillips JD (2001) Designing matched bandpass and azimuthal filters for the separation of potential-field anomalies by source region and source type. In: *15th Geophysical Conference and Exhibition, Australian Society of Exploration Geophysicists, Expanded Abstracts, CD-ROM*.

Phillips JD (2005) Can we estimate total magnetization directions from aeromagnetic data using Helbig's integrals? *Earth, Planets and Space* 57: 681–689.

Phillips JD, Nabighian MN, Smith DV, and Li Yaoguo (2007) Estimating locations and total magnetization vectors of compact magnetic sources from scalar, vector, or tensor magnetic measurements through combined Helbig and Euler analysis. In: *Society of Exploration Geophysicists, Seventy-Seventh Annual Meeting, San Antonio, Texas, September 23–28*.

Pilkington M (1997) 3-D magnetic imaging using conjugate gradients. *Geophysics* 62: 1132–1142.

Pilkington M and Beiki M (2013) Mitigating remanent magnetization effects in magnetic data using the normalized source strength. *Geophysics* 78(3): J25–J32.

Pilkington M and Crossley DJ (1986) Determination of crustal interface topography from potential fields. *Geophysics* 51: 1277–1284.

Portniaguine O and Zhdanov M (2002) 3D magnetic inversion with data compression and image focusing. *Geophysics* 67: 1532–1541.

Rao DB and Babu NR (1991) A rapid methods for three-dimensional modeling of magnetic anomalies. *Geophysics* 56: 1729–1737.

Rasmussen R and Pedersen LB (1979) End corrections in potential field modeling. *Geophysical Prospecting* 27: 749–760.

Reford S (2006) Gradient enhancement of the total magnetic field. *The Leading Edge* 25: 59–66.

Reid AB, Allsop JM, Granser H, Millett AJ, and Somerton IW (1990) Magnetic interpretation in three dimensions using Euler deconvolution. *Geophysics* 55: 80–91.

Reynolds JM (1997) *An Introduction to Applied and Environmental Geophysics*. John Wiley & Sons.

Ridsdill-Smith TA (1998a) Separating aeromagnetic anomalies using wavelet matched filters. In: *68th Annual International Meeting, SEG, Expanded Abstracts*, pp. 550–553.

Ridsdill-Smith TA (1998b) Separation filtering of aeromagnetic data using filter banks. *Exploration Geophysics* 29: 577–583.

Ridsdill-Smith TA and Dentith MC (1999) The wavelet transform in aeromagnetic processing. *Geophysics* 64: 1003–1013.

Roest W and Pilkington M (1993) Identifying remanent magnetization effects in magnetic data. *Geophysics* 58: 653–659.

Roest WR, Verhoef J, and Pilkington M (1992) Magnetic interpretation using the 3D analytic signal. *Geophysics* 57: 116–125.

Salem A, Williams S, Fairhead D, Smith R, and Dhananjay Ravat D (2008) Interpretation of magnetic data using tilt-angle derivatives. *Geophysics* 73(1): L1–L10.

Salem A, Lei K, Green C, Fairhead J, and Stanley G (2010) Removal of cultural noise from high-resolution aeromagnetic data using a two stage equivalent source approach. *Exploration Geophysics* 41(2): 163–169.

Saltus RW and Blakely RJ (1983) HYPERMAG, an interactive, two-dimensional gravity and magnetic modeling program. US Geological Survey Open File Report, pp 83–241.

Saltus RW and Blakely RJ (1993) HYPERMAG, an interactive, 2- and 221/2-dimensional gravity and magnetic modeling program, version 3.5. US Geological Survey Open File Report, pp 93–287.

Sanchez V, Li Y, Nabighian M, and Wright D (2008) Numerical modeling to identify relative contributions of high-order magnetic moments in UXO discrimination. *IEEE Transactions on Geoscience and Remote Sensing* 46: 2568–2583.

Santos M and Li Y (2012) Application of 3D magnetic amplitude inversion to Fe oxide–Cu–Au deposits at low magnetic latitudes: A case study from Carajas Mineral Province, Brazil. In: *82nd SEG Annual Meeting*.

Schmidt PW and Clark DA (1997) Directions of magnetization and vector anomalies derived from total field surveys. *Preview* 70: 30–32.

Schmidt PW and Clark DA (1998) The calculation of magnetic components and moments from TMI: A case history from the Tuckers igneous complex, Queensland. *Exploration Geophysics* 29: 609–614.

Sharma PV (1966) Rapid computation of magnetic anomalies and demagnetization effects caused by bodies of arbitrary shape. *Pure and Applied Geophysics* 64: 89–109.

Sharma PV (1968) Short note: Demagnetization effect of a rectangular prism. *Geophysics* 33: 132–134.

Shearer SE (2005) *3-D inversion of magnetic data in the presence of remanent Magnetization*. MS Thesis, Colorado School of Mines.

Sheriff RE (1989) *Geophysical Methods*. New Jersey: Prentice Hall.

Shuey RT and Pasquale AS (1973) End corrections in magnetic profile interpretation. *Geophysics* 38: 507–512.

Silva JBC (1986) Reduction to the pole as an inverse problem and its application to low-latitude anomalies. *Geophysics* 51: 369–382.

Silva JBC, Gerald W, and Hohmann GW (1983) Nonlinear magnetic inversion using a random search method. *Geophysics* 48: 1645–1658.

Silva JBC and Hohmann GW (1981) Interpretation of three-component borehole magnetometer data. *Geophysics* 46: 1721–1731.

Smith RS, Thurston JG, and MacLeod IN (1998) ISPITM—The improved source parameter imaging method. *Geophysical Prospecting* 46: 141–151.

Spector A (1968) *Spectral Analysis of Aeromagnetic Data*. PhD Thesis, University of Toronto.

Spector A and Bhattacharyya BK (1966) Energy density spectrum and autocorrelation function of anomalies due to simple magnetic models. *Geophysical Prospecting* 14: 242–272.

Spector A and Grant F (1970) Statistical models for interpreting aeromagnetic data. *Geophysics* 35: 293–302.

Stavrev P and Gerovska D (2000) Magnetic field transforms with low sensitivity to the direction of source magnetization and high centricity. *Geophysical Prospecting* 48: 317–340.

Syberg FJR (1972) A Fourier method for the regional-residual problem of potential fields. *Geophysical Prospecting* 20: 47–75.

Talwani M (1965) Computation with the help of a digital computer of magnetic anomalies caused by bodies of arbitrary shape. *Geophysics* 30: 797–817.

Talwani M and Heirtzler JR (1964) Computation of magnetic anomalies caused by two-dimensional structures of arbitrary shape. In: *Computers in the Mineral Industries*. Stanford University Publications of the Geological Sciences.

Telford WM and Sheriff RE (1990) *Applied Geophysics*. Cambridge University Press.

Telford WM, Geldart LP, and Sheriff RE (1996) *Applied Geophysics*, 3rd ed. Cambridge: Cambridge University Press.

Thompson DT (1982) EULDPH – A new technique for making computer-assisted depth estimates from magnetic data. *Geophysics* 47: 31–37.

Thurston JB and Smith RS (1997) Automatic conversion of magnetic data to depth, dip, and susceptibility contrast using the SPITM method. *Geophysics* 62: 807–813.

Tikhonov AN and Arsenin VY (1979) *Methods for Solving Ill-Posed Problems*. Moscow: Nauka.

Traynin P and Hansen RO (1993) Magnetic modeling for highly permeable bodies with remanent magnetization. In: 63rd Annual International Meeting, Society of Exploration Geophysicists.

Tricomi FG (1985) *Integral Equations*. Dover Publication, 254 pp.

Vacquier V, Steenland NC, Henderson RG, and Zietz I (1951) Interpretation of aeromagnetic maps. *Memoir 47*. Geological Society of America.

Wahba G (1990) *Spline Models for Observational Data: CBMS-NSF, Regional CONFERENCE SERIES in APPLIED Mathematics*. Philadelphia, PA: SIAM, 211–224.

Wallace Y (2007) 3D modelling of banded iron formation incorporating demagnetization – A case study at the Musselwhite Mine, Ontario, Canada. *Exploration Geophysics* 38: 254–259.

Wang B (1995) Effective approaches to handling non-uniform data coverage problem for wide-aperture refraction/reflection profiling. In: *65th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts*, pp. 659–662.

Wang M, Di Q, Xu K, and Wang R (2004) Magnetization vector inversion equations and forword [*sic*] and inverse [*sic*] 2-d model study. *Chinese Journal of Geophysics* 47: 601–609.

Wang X and Hansen RO (1990) Inversion for magnetic anomalies of arbitrary three-dimensional bodies. *Geophysics* 55: 1321–1326.

Watts A (1997) Exploration for nickel in the 90s, or 'Till depth us do part'. In: *Exploration 97*.

Webring M (1985) SAKI, A Fortran program for generalized linear inversion of gravity and magnetic profiles. US Geological Survey Open File Report 85-122.

Werner S (1955) Interpretation of magnetic anomalies of sheet-like bodies. *Sveriges Geologiska Undersökning, Series C: Arsbok* 43, No. 6.

Whitehill D (1973) Automated interpretation of magnetic anomalies using the vertical prism model. *Geophysics* 38: 1070–1087.

Wright SJ (1997) *Primal-Dual Interior-Point Methods*. Philadelphia, PA: SIAM.

Zeyen H and Pous J (1991) A new 3-D inversion algorithm for magnetic total field anomalies. *Geophysical Journal International* 104(Issue 3): 583–591.

Zhang C, Mushayandebvu MF, Reid AB, Fairhead JD, and Odegard ME (2000) Euler deconvolution of gravity tensor gradient data. *Geophysics* 65: 512–520.

Zietz I and Andreasen GE (1967) Remanent magnetization and aeromagnetic interpretation. *Mining Geophysics* 2: 569–590.

This page intentionally left blank

11.12 Tools and Techniques: Gravitational Method

JD Phillips, U.S. Geological Survey, Denver, CO, USA

Published by Elsevier B.V.

11.12.1	**Introduction: The Role of the Gravitational Method in Near-Surface Investigations**	394
11.12.2	**Fundamentals**	394
11.12.2.1	Newton's Law of Gravitation	394
11.12.2.2	The Gravity Field	394
11.12.2.3	The Gravitational Potential	395
11.12.2.4	Units of Gravity	395
11.12.3	**Gravity Fields of Simple Sources in Cartesian Coordinates**	395
11.12.3.1	3-D Sources	395
11.12.3.2	2-D Sources	396
11.12.3.3	1-D Sources	396
11.12.4	**Gravity Gradient Tensor Components**	396
11.12.4.1	Units of Gravity Gradients	396
11.12.5	**Poisson's Relation**	396
11.12.6	**The Earth's Gravity Field**	396
11.12.7	**Density**	397
11.12.7.1	Measuring Density	397
11.12.7.2	Densities of Geologic Materials	397
11.12.8	**Instrumentation**	397
11.12.8.1	Relative Instruments	397
11.12.8.2	Absolute Instruments	398
11.12.8.3	Gradiometers	398
11.12.8.4	Satellites	399
11.12.9	**Field Procedures on Land**	399
11.12.10	**Data Reduction**	400
11.12.10.1	Free-Air Anomalies	401
11.12.10.2	Bouguer Anomalies	402
11.12.10.2.1	Curvature correction	403
11.12.10.2.2	Terrain correction	403
11.12.10.3	Isostatic Residual Anomalies	404
11.12.10.4	Eötvös Correction	404
11.12.11	**Gravity Profiles and Grids**	404
11.12.12	**Filtering**	405
11.12.13	**Interpretation: Forward Modeling**	405
11.12.13.1	2.5-D Polygonal Bodies	405
11.12.13.2	Laminar Bodies	405
11.12.13.3	Gridded Surfaces	405
11.12.13.4	Faceted Bodies	406
11.12.13.5	Voxel Models	407
11.12.14	**Interpretation: Source Location Methods Applied to Gravity**	407
11.12.14.1	Werner Deconvolution	407
11.12.14.2	Euler Deconvolution	408
11.12.15	**Interpretation: Inversion**	409
11.12.15.1	Trial and Error	409
11.12.15.2	Iterative Inversions for Density Interfaces	409
11.12.15.2.1	Basin thickness from gravity profile data: Bott's method	409
11.12.15.2.2	Density interface from gravity profile data: The Parker–Oldenburg method	409
11.12.15.2.3	Density interfaces from gridded gravity: The Cordell–Henderson method	409
11.12.15.2.4	Basin/basement gravity separation and basin thickness: The Jachens–Moring method	409
11.12.15.3	Physical Property Inversions	410
11.12.16	**Field Examples**	410
11.12.16.1	Estimating Overburden Thickness and Bedrock Depth at the Tooele Army Depot, Utah, Using Constrained Gravity Inversion	410
11.12.16.2	Estimating Thickness of Sedimentary Rocks in the Yukon Flats Basin, East-Central Alaska, Using Constrained Iterative Gravity Inversion	411

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00197-4

11.12.16.3 Estimating Terrain Density from Gravity Profile Data, National Petroleum Reserve Alaska 411
11.12.16.4 Using Time-Lapse Gravity to Monitor Subsurface Water Storage in a Former Coal Mine, Arvada, Colorado 413
11.12.17 Summary 414
Disclaimer 417
References 417

11.12.1 Introduction: The Role of the Gravitational Method in Near-Surface Investigations

The gravitational (or gravity) method is used to investigate density variations within the subsurface. The density variations of interest may occur at depths of several meters to tens of meters, as in depth-to-bedrock investigations, or at depths of several kilometers, as in sedimentary basin thickness investigations. The gravitational method is often used in conjunction with other geophysical methods, such as the magnetic method (Chapter 11.11) or the seismic method (Chapter 11.06), which target similar physical properties at similar depths.

Measurements of the gravity field have traditionally been made at point locations on the Earth's surface, in boreholes, or along ship tracks at sea. In recent decades, the shipborne technology has been adapted to aircraft to allow airborne gravity surveys, which can be used for geologic mapping under cover. New instruments can measure gradients of the gravity field from moving platforms such as aircraft, offering data of higher resolution and shorter wavelengths. Longer wavelength components of the gravity field can be measured by tracking of artificial satellites.

Gravity measurements repeated over time can be used to monitor changes in the density of the subsurface due to hydrocarbon extraction, aquifer withdrawal and recharge, or earthquake displacement. In areas such as the Gulf of Mexico, the gravity method is used for mapping of salt horizons, which can be difficult to image in seismic reflection data.

For more details on the history of the gravity method as applied to exploration, along with some additional field examples, see Nabighian et al. (2005).

11.12.2 Fundamentals

11.12.2.1 Newton's Law of Gravitation

The gravitational method is based on Newton's law of gravitation (Newton, 1687), which states that the attractive force F between two point masses, m_1 and m_2, is proportional to the product of the masses and inversely proportional to the square of the distance r between them:

$$F = \gamma \frac{m_1 m_2}{r^2} \quad [1]$$

The fundamental constant of proportionality, γ (or G), is called the gravitational constant and is determined experimentally. Expressed in SI units, the force is in newtons (N), mass is in kilograms (kg), and distance is in meters (m). The gravitational constant is known to within three or four significant digits and has an approximate value of

$$\gamma = 6.674 \times 10^{-11}\ \mathrm{N m^2 kg^{-2}} = 6.674 \times 10^{-11}\ \mathrm{m^3 kg^{-1} s^{-2}} \quad [2]$$

The constant is difficult to measure accurately, and its recommended value has changed over the years based on updated experimental results (Table 1).

Equation [1] also holds for nonpoint masses, in which case r is the distance between the two centers of mass.

11.12.2.2 The Gravity Field

The gravitational force can be thought of as a vector quantity directed along the axis between the two masses. If one mass is removed or regarded as a test mass of unit magnitude, the result is an acceleration that represents the gravity field

$$\mathbf{g}(\mathbf{r}) = -\gamma \frac{m_1}{|\mathbf{r}|^2}\hat{\mathbf{r}} \quad [3]$$

Using the convention of Kellogg (1953) and Blakely (1995), $\hat{\mathbf{r}}$ is a unit vector directed from the center of mass of m_1 toward the observation point. For terrestrial applications involving absolute gravity measurements, m_1 represents the mass of the Earth. For applications involving relative gravity measurements, m_1 represents an element of the anomalous subsurface mass. Through the principle of superposition, the gravity field of a collection of masses is the vector sum of the gravity fields of the individual masses. Similarly, the gravity field produced by an arbitrary distribution of density $\rho(\mathbf{r})$ is given by the volume integral

$$\mathbf{g}(\mathbf{r}) = -\gamma \int_v \frac{\rho(\mathbf{r})}{|\mathbf{r}|^2}\hat{\mathbf{r}}\, dv \quad [4]$$

The vertical component of the gravity field $g_z(\mathbf{r})$ is the quantity measured by most gravimeters and used in the interpretation of subsurface density variations. In Cartesian coordinates, the vertical gravity field outside of a point mass, sphere of uniform density, or spherical shell of uniform density having total mass m and a center at (x_0, y_0, z_0) is given by

$$g_z(x, y, z) = \frac{-\gamma m (z - z_0)}{\left((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\right)^{3/2}} \quad [5]$$

Additional equations for a spherical source are given in Table 2.

Table 1 CODATA recommended values of the gravitational constant (Committee on Data for Science and Technology, International Council for Science)

Year	*Recommended value of the gravitational constant ($\times 10^{-11}$ m^3 kg^{-1} s^{-2})*
1986	6.67259 ± 0.00085
1998	6.673 ± 0.010
2002	6.6742 ± 0.0010
2006	6.67428 ± 0.00067
2010	6.67384 ± 0.00080

Table 2 Equations for the gravitational potential, gravity anomaly components, and gradient tensor components for a point mass, uniform-density sphere, or uniform-density spherical shell. The source is centered at (x_0, y_0, z_0) and has mass m. Note that the z-axis is positive down, so ($z-z_0$) is negative for observations above the source

Component	*Anomaly*
Gravitational potential	$U(x,y,z)=\dfrac{\gamma m}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{1/2}}$
North component of gravity	$g_x(x,y,z)=\dfrac{-\gamma m(x-x_0)}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{3/2}}$
East component of gravity	$g_y(x,y,z)=\dfrac{-\gamma m(y-y_0)}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{3/2}}$
Vertical gravity	$g_z(x,y,z)=\dfrac{-\gamma m(z-z_0)}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{3/2}}$
G_{xx} tensor component	$G_{xx}(x,y,z)=\gamma m\dfrac{2(x-x_0)^2-(y-y_0)^2-(z-z_0)^2}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{5/2}}$
G_{yy} tensor component	$G_{yy}(x,y,z)=\gamma m\dfrac{2(y-y_0)^2-(x-x_0)^2-(z-z_0)^2}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{5/2}}$
G_{zz} tensor component	$G_{zz}(x,y,z)=\gamma m\dfrac{2(z-z_0)^2-(x-x_0)^2-(y-y_0)^2}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{5/2}}$
G_{xy} or G_{yx} tensor component	$G_{xy}(x,y,z)=\dfrac{-3\gamma m(x-x_0)(y-y_0)}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{5/2}}$
G_{xz} or G_{zx} tensor component	$G_{xz}(x,y,z)=\dfrac{-3\gamma m(x-x_0)(z-z_0)}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{5/2}}$
G_{yz} or G_{zy} tensor component	$G_{yz}(x,y,z)=\dfrac{-3\gamma m(y-y_0)(z-z_0)}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{5/2}}$

11.12.2.3 The Gravitational Potential

The gravity field is a vector quantity and can be expressed as the gradient of a scalar potential U:

$$\mathbf{g}(\mathbf{r})=\nabla U(\mathbf{r}) \qquad [6]$$

For this reason, a gravity field, like a magnetic field, is classified as a potential field. The gravity potential (1) satisfies Laplace's equation $\nabla^2 U=0$ outside of source regions and (2) satisfies Poisson's equation $\nabla^2 U=-4\pi\gamma\rho$ inside of source regions, where ρ represents the density.

Because the gravitational potential satisfies Laplace's equation, there can be no concentrations of potential within a source-free region. All minima or maxima of the potential must occur on the boundary between the source-free region and the region containing the sources. For terrestrial applications, this boundary is the surface of the Earth.

The gravitational potential in a source-free region fits the definition of a harmonic function – it satisfies Laplace's equation, it has continuous first derivatives, and its second derivatives exist. Because any spatial derivative of a harmonic function is also a harmonic function, any component of the gravity field will also satisfy Laplace's equation in a source-free region. This becomes important in the discussion of gravity gradient tensors.

The gravitational potential and its spatial derivatives are also homogeneous functions, in the sense that they satisfy Euler's homogeneity equation

$$x\frac{\partial U}{\partial x}+y\frac{\partial U}{\partial y}+z\frac{\partial U}{\partial z}=nU \qquad [7]$$

Here, n is an integer called the degree of homogeneity. This property allows interpretation of density source locations using Euler deconvolution.

A final property of the gravity field is provided by Gauss' law: The total mass M in a region is proportional to the normal component of gravitational attraction integrated over the closed boundary S of the region:

$$\int_S \mathbf{g}\cdot\hat{\mathbf{n}}\,dS=-4\pi\gamma M \qquad [8]$$

From a practical standpoint, this law allows us to estimate the total anomalous mass underlying an area of interest by integrating the vertical gravity field over a horizontal plane defining the area:

$$M=\frac{1}{2\pi\gamma}\iint g_z\,dx\,dy \qquad [9]$$

For best results, the integration should be performed on a closed positive residual gravity field after a regional gravity field has been removed.

11.12.2.4 Units of Gravity

The standard unit of gravity is the gal (or galileo); the unit name is capitalized (Gal):

$$1\,\text{Gal}=1\,\text{cm}\,\text{s}^{-2}=0.01\,\text{m}\,\text{s}^{-2}(\text{SI}) \qquad [10]$$

Standard gravity at sea level varies from 978 Gal at the Earth's equator to 983 Gal at the Earth's poles.

Gravity anomalies produced by geological sources are typically expressed in units of milligals (mGal):

$$1\,\text{mGal}=10^{-3}\text{Gal}=10^{-5}\text{m}\,\text{s}^{-2}(\text{SI}) \qquad [11]$$

Mountains, sedimentary basins, and smaller geological sources produce gravity anomalies of tens to hundreds of milligals.

Microgravity anomalies produced by variations in soil density, by archeological sources, or by time-lapse changes in subsurface oil or water are typically expressed in units of microgals (μGal):

$$1\,\mu\text{Gal}=10^{-3}\text{mGal}=10^{-6}\text{Gal}=10^{-6}\text{m}\,\text{s}^{-2}(\text{SI}) \qquad [12]$$

11.12.3 Gravity Fields of Simple Sources in Cartesian Coordinates

11.12.3.1 3-D Sources

The vertical gravity field outside of a point mass, uniform sphere, or uniform spherical shell is given by eqn [5]. By

integrating eqn [5] along the z-axis, we derive the vertical gravity field outside of a semi-infinite vertical line source or a uniform-density vertical cylinder with a hemispherical cap:

$$g_z(x,y,z)=\frac{\gamma m}{\left((x-x_0)^2+(y-y_0)^2+(z-z_0)^2\right)^{1/2}} \quad [13]$$

Here, m is the mass per unit length.

11.12.3.2 2-D Sources

By integrating eqn [5] along the y-axis, we derive the vertical gravity field produced by an infinite horizontal line of mass or an infinite horizontal cylinder or cylindrical shell of uniform density:

$$g_z(x,z)=\frac{-2\gamma m(z-z_0)}{(x-x_0)^2+(z-z_0)^2} \quad [14]$$

By integrating eqn [14] along the x- and z-axes, we derive the vertical gravity field of a dipping semi-infinite horizontal sheet:

$$g_z(x,z)=-2\gamma m\left(\frac{1}{2}\ln\left((x-x_0)^2+(z-z_0)^2\right)\sin d + \arctan\left(\frac{x-x_0}{z-z_0}\right)\cos d\right) \quad [15]$$

In this case, m is the mass per unit area and d is the dip as measured from the positive x-axis.

11.12.3.3 1-D Sources

The vertical gravity field above an infinite horizontal layer of thickness t and uniform density ρ is a constant given by the Bouguer slab formula:

$$g_z(x,y,z)=2\pi\gamma\rho t \quad [16]$$

Below the layer, the vertical gravity is the negative of [16], and within the layer, the vertical gravity decreases linearly with depth.

11.12.4 Gravity Gradient Tensor Components

As mentioned earlier, some newer instruments are capable of measuring gravity gradients, which are normally presented as tensor components. The gravity gradient tensor contains nine components, five of which are independent:

$$\begin{vmatrix} G_{xx} & G_{xy} & G_{xz} \\ G_{yx} & G_{yy} & G_{yz} \\ G_{zx} & G_{zy} & G_{zz} \end{vmatrix} = \begin{vmatrix} \frac{\partial g_x}{\partial x} & \frac{\partial g_x}{\partial y} & \frac{\partial g_x}{\partial z} \\ \frac{\partial g_y}{\partial x} & \frac{\partial g_y}{\partial y} & \frac{\partial g_y}{\partial z} \\ \frac{\partial g_z}{\partial x} & \frac{\partial g_z}{\partial y} & \frac{\partial g_z}{\partial z} \end{vmatrix} = \begin{vmatrix} \frac{\partial^2 U}{\partial x^2} & \frac{\partial^2 U}{\partial x\partial y} & \frac{\partial^2 U}{\partial x\partial z} \\ \frac{\partial^2 U}{\partial y\partial x} & \frac{\partial^2 U}{\partial y^2} & \frac{\partial^2 U}{\partial y\partial z} \\ \frac{\partial^2 U}{\partial z\partial x} & \frac{\partial^2 U}{\partial z\partial y} & \frac{\partial^2 U}{\partial z^2} \end{vmatrix} \quad [17]$$

Because differentiation is commutative, it is clear that $G_{xy}=G_{yx}$, $G_{xz}=G_{zx}$, and $G_{yz}=G_{zy}$. In addition, Laplace's equation requires that $G_{zz}=G_{xx}+G_{yy}$. By convention, x is north, y is east, and z is down.

11.12.4.1 Units of Gravity Gradients

The standard unit for gravity gradients is the eötvös (E):

$$1\,\mathrm{E}=10^{-9}\mathrm{Gal\,cm^{-1}}=10^{-7}\mathrm{Gal\,m^{-1}}=10^{-4}\mathrm{mGal\,m^{-1}} =10^{-9}\mathrm{s^{-2}(SI\,or\,cgs)} \quad [18]$$

The near-spherical Earth produces a vertical gravity gradient of $G_{zz}=3080\mathrm{E}$ ($=0.308$ mgal m$^{-1}=3.08$ μGal cm^{-1}) at the Earth's surface and horizontal gravity gradients of $G_{xx}=G_{yy}=-1540\mathrm{E}$. These values can be derived from equations in Table 2 using the mean radius (6371 km) and mass (5.972×10^{24} kg) of the Earth. The factor-of-two and sign differences result from Laplace's equation.

11.12.5 Poisson's Relation

Poisson's relation states that, for a body with uniform density contrast and magnetization, the component of gravitational attraction parallel to the magnetization is proportional to the magnetic scalar potential of the body. This means that if we know the density contrast $\Delta\rho$ and the magnetization m_z of a body with vertical magnetizaton, we can calculate its magnetic field from its vertical gravity field g_z and vice versa. For such a body, the magnetic potential can be written as

$$V=-(C_m m_z/\gamma\Delta\rho)g_z \quad [19]$$

where γ is the universal gravitational constant.

If the gravity (or magnetic) field of a source is known in closed form, Poisson's relation can be used to obtain its magnetic (or gravity) field in closed form. The relation is also useful for transforming gravity fields to pseudo-magnetic fields and magnetic fields to pseudo-gravity fields using Fourier filters. This is discussed further in Section 11.12.12.

11.12.6 The Earth's Gravity Field

Gravity measurements are dominated by the overall gravity field of the Earth. This field must be removed from the measurements in order to study the gravity field of the local anomalous mass. The total mass M of the Earth is a fixed quantity, but because the Earth is an oblate spheroid, the distance from the Earth's surface to the center of mass varies with latitude. Because the Earth is rotating, there is a latitude-dependent acceleration that must be removed. Finally, because the Earth's surface is not flat, there is an elevation effect.

For the purposes of gravity reduction, the Earth is treated as a rotating oblate spheroid having an equatorial radius of $a=6378.137$ km and a flattening of $f=1/298.257$, where flattening is the difference between the equatorial and polar radii divided by the equatorial radius. These numbers are taken from the latest (1980) version of three geodetic reference systems sanctioned by the International Association of Geodesy (Blakely, 1995). The surface of the oblate spheroid is called the 'reference ellipsoid.' The gravitational potential U of the spheroid is the sum of its self-gravitational potential U_g, which is completely determined by M, f, and a, and its rotational

potential U_r, which is determined by the angular velocity ω, the latitude λ, and the ellipsoidal radius.

Based on the 1980 geodetic reference system, the 'theoretical gravity' or 'normal gravity' on the reference ellipsoid at latitude λ is given by Blakely (1995):

$$g_0 = -\frac{\partial U}{\partial r} = g_e\left(\frac{1+k\sin^2\lambda}{\sqrt{1-e^2\sin^2\lambda}}\right) = 9.7803267714\left(\frac{1+0.00193185138639\sin^2\lambda}{\sqrt{1-0.00669437999013\sin^2\lambda}}\right) \quad [20]$$

Here, g_e is the gravitational acceleration at the equator in m s^{-2} and k and e are constants that depend only on M, f, a, and ω. In some cases (LaFehr, 1991a; LaFehr and Nabighian, 2012), an earlier (1967) approximation to [20] is used:

$$g_0 = 9.78031846\left(1+0.0053024\sin^2\lambda - 0.0000058\sin^2 2\lambda\right) \quad [21]$$

The reference ellipsoid is equivalent to an equipotential surface at sea level produced by a rotating, uniform-density Earth. The actual equipotential surface corresponding to mean sea level is called the 'geoid.' Elevations (which are used in gravity reduction) are measured relative to the geoid, so there is an inconsistency in the overall approach. Someday, the reference ellipsoid may be replaced by the actual geoid, but right now, for purposes of crustal gravity studies, geophysicists ignore the inconsistency and treat the theoretical gravity on the ellipsoid as if it were the theoretical gravity on the geoid.

11.12.7 Density

The physical property associated with the gravitational method is density or density contrast. Geophysicists measure density in kilograms per cubic meter (kg m^{-3}, SI units) or grams per cubic centimeter (g cm^{-3}, cgs units). The mean density of the Earth is 5.515 g cm^{-3}, but most of the mass is concentrated in the core and mantle. In the core, estimated densities vary from 13.0 or 13.1 g cm^{-3} at the Earth's center to 9.9 g cm^{-3} at the outer core boundary. In the mantle, estimated densities vary from 5.7 or 5.6 g cm^{-3} in the deep mantle to 3.3 or 3.4 g cm^{-3} at the top of the mantle. The densities of near-surface Earth materials vary over a relatively small range from values near 1.0 g cm^{-3} for water or ice to values of around 4.0 g cm^{-3} for the highest density rocks. Air, with a density of 0.001 g cm^{-3}, is generally treated as having zero density for purposes of geophysical modeling.

11.12.7.1 Measuring Density

Because rocks are heterogeneous collections of minerals, pore space or fractures, and pore fluids, their densities are more properly referred to as bulk densities. In a laboratory, the bulk density of a rock is measured as the mass of the rock divided by the volume of the rock. The mass is measured on a scale or balance, and the volume is measured by the displacement of water or oil (Archimedes' principle).

Bulk density can be measured in a borehole using a density (gamma–gamma) logging tool, which uses a radioactive source to emit gamma rays into the adjacent rock and uses one or more detectors at fixed distances from the source to count the gamma rays returned as a result of Compton scattering. The gamma ray count is related to the electron density of the rock, which in turn is related to the bulk density.

Gravity gradiometry can also be used to measure the bulk densities of horizontal density layers within a borehole. Based on the discussion of the Bouguer slab formula in Section 11.12.3.3, the difference between two borehole gravity measurements separated by a vertical distance Δt within a horizontal layer of thickness $t_1 \geq \Delta t$ and density ρ_1 is

$$\Delta g_z = 2\pi\gamma\rho_1\Delta t \quad [22]$$

Thus, the bulk density of the layer is

$$\rho_1 = \frac{\Delta g_z}{2\pi\gamma\Delta t} \quad [23]$$

Sedimentary rocks can contain pore space that is typically filled with air or water and less typically with natural gas or crude oil. These pore spaces and pore fluids contribute to the bulk density of the rock. At depth, rocks containing pore space or fractures are compressed by the weight of the overburden, which reduces the volume of the pore space and thus increases the bulk density. The increase of density with depth in a sedimentary basin is of interest to exploration geophysicists, primarily because it affects the seismic velocity at depth and thus the conversion of seismic time sections to depth sections.

11.12.7.2 Densities of Geologic Materials

Table 3 and Figure 1 present densities of common Earth materials.

11.12.8 Instrumentation

Although a great many instruments have been invented to measure gravity, only a few are currently in use. These include relative gravimeters, which measure the relative change in the gravity field from place to place; absolute gravity meters, which measure the absolute value of gravity at a station; gradiometers, which measure gravity gradients; and satellites.

11.12.8.1 Relative Instruments

Relative gravimeters have achieved a level of ruggedness and accuracy that make them the standard instrument for rapid field measurements on the ground, at sea, in boreholes, and in aircraft. These instruments generally use weights, springs, beams, and mirrors to measure small variations in the vertical component of gravity through displacement of a responsive element from an equilibrium position. Because the displacements are small, they must be magnified using optical, mechanical, or electrical means. The LaCoste–Romberg gravimeter, in use since 1939, uses a mechanical 'zero-length' spring system, and the newer Scintrex CG-5 gravimeter uses a fused quartz spring system with electrostatic nulling. Modern instruments can incorporate automatic leveling, precision timekeeping, and built-in GPS sensors. Precisions are in the 2 μGal range, with repeatability in the 5 μGal range.

Table 3 Densities of Earth materials. Sources include Dobrin (1960) and various websites

	Reported range of densities (g cm^{-3})	*Reported (or estimated) average densities (g cm^{-3})*
Earth material		
Air	0.0012041–0.001293	0.001275
Glacial ice	0.88–0.92	0.9
Ice	0.917	0.917
Water	1	1
Seawater	1.025	1.025
Hydrocarbons		
Crude oil	0.79–0.97	0.9
Coal, bituminous	1.1–1.5	(1.2)
Coal, anthracite	1.2–1.8	1.3
Oil shale	2.0–2.2	(2.1)
Unconsolidated sediment		
Loam	1.28–1.44	(1.36)
Loess	1.01–1.93	1.64
Soil	1.1–2.0	1.92
Silt	1.36–2.20	1.93
Recent alluvium	1.53–2.0	1.98
Sand	1.40–2.30	2
Earth	1.2–2.4	2
Gravel	1.7–2.4	2
Evaporites		
Rock salt	2.1–2.6	2.22
Gypsum	2.2–3.3	2.3
Anhydrite	2.90–2.96	(2.93)
Sedimentary rocks		
Diatomite	0.45–1.18	(0.815)
Chalk	1.53–2.60	2.01
Clay	1.2–2.6	2.21
Marl	2.0–2.6	(2.3)
Sandstone	1.61–2.8	2.32, 2.35, 2.50
Shale	1.57–3.20	2.40, 2.42, 2.45
Graywacke	2.4–2.5	(2.45)
Limestone	1.55–2.9	2.50, 2.54, 2.55
Siltstone	2.3–2.8	(2.55)
Conglomerate	2.6–2.7	(2.65)
Argillite	2.5–2.8	(2.65)
Dolomite	2.28–2.90	2.7, 2.8
Chert	2.0–3.6	(2.8)
Metamorphic rocks		
Marble	2.1–2.9	(2.5)
Schist	2.39–3.19	2.64, 2.70
Quartzite	2.0–2.8	2.7
Phyllite	2.2–3.2	(2.7)
Slate	2.5–2.9	2.79
Gneiss	2.59–3.00	2.8
Amphibolite	2.79–3.14	2.96, 3.00
Greenstone	3.0–3.1	(3.05)
Eclogite	3.20–3.54	3.37
Igneous rocks		
Rhyolite	2.1–2.7	2.52
Andesite	2.4–2.8	2.61
Granite	2.5–2.9	2.64, 2.70
Granodiorite	2.668–2.8	(2.734)
Syenite	2.630–2.899	(2.7645)
Anorthosite	2.64–2.92	(2.78)
Diabase	2.5–3.2	(2.85)
Norite	2.72–3.02	(2.87)
Diorite	2.721–3.0	2.9
Basalt	2.7–3.3	2.90, 2.99
Gabbro	2.7–3.5	3.03
Pyroxenite	3.10–3.318	3.209
Dunite	3.289	3.289
Peridotite	3.1–3.4	3.37
Metallic ore		
Copper ore	2	2
Chromite	4.5–4.8	(4.65)
Iron ore	4.2–5.3	(4.75)
Lead ore	7.5	7.5

(*Continued*)

Table 3 (Continued)

11.12.8.2 Absolute Instruments

The absolute value of gravity at a site can be measured using a pendulum or using laser interferometric measurements of a falling mass. Pendulum measurements are based on the relationship between the period of the pendulum and the acceleration of gravity. Carefully controlled conditions are required in order to obtain high accuracy measurements, which limits the use of pendulum instruments in the field.

Absolute gravity meters based on repeated measurements of a falling mass have been deployed as field instruments at ground-based sites. The model A10 from Micro-g LaCoste is one example; accuracies of 10 μGal are claimed.

11.12.8.3 Gradiometers

Instruments that measure gradients of the gravity field include the historical torsion balance and modern instruments designed for deployment in aircraft. The torsion balance was used up until the mid-1930s to conduct gravity surveys on land for structures related to petroleum. The use of the balance was labor-intensive, requiring clearing and leveling of the adjacent land and long setup and wait times. Measurements could not be made at all in rugged terrain (Bell and Hansen, 1998).

Modern gravity gradiometers use pairs of accelerometers mounted on rotating disks to measure gradients of the gravity field. The Falcon system (Dransfield and Lee, 2004; Lee, 2001) uses a single disk containing four pairs of accelerometers, and rotating about a nearly vertical axis, to measure two components of the gravity tensor:

$$G_{xy} \text{ and } \left(G_{xx} - G_{yy}\right)/2 \qquad [24]$$

These two components, combined as a complex signal, can be used to generate all other components from Laplace's equation in the Fourier domain.

The Bell Air-FTG (full tensor gravity) system (Murphy, 2004) uses three disks, each containing two pairs of accelerometers. The disks rotate along orthogonal axes, each oriented at the same angle to the vertical. All components of the gravity tensor are measured.

Both types of instrument are mounted on stable platforms that maintain their orientation relative to the vertical. Corrections must be made for aircraft accelerations and for the effect of the mass of the aircraft as it rotates around the instrument

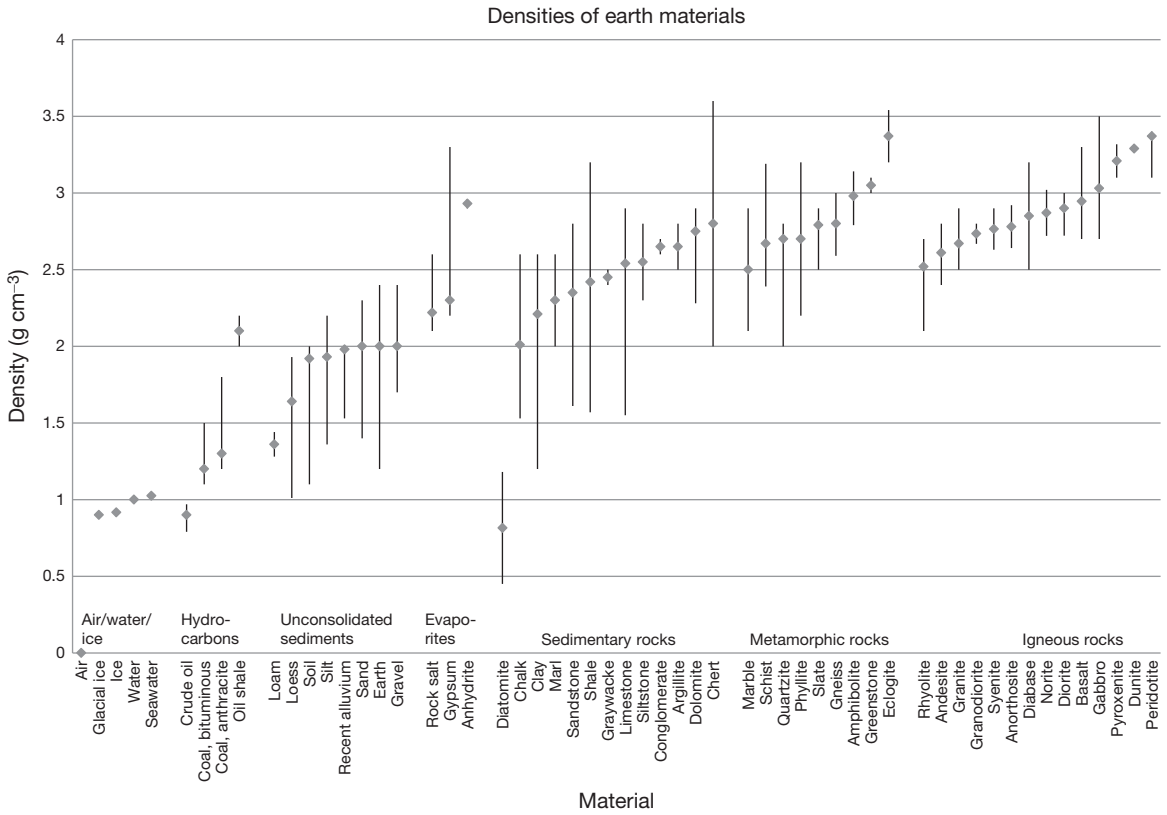


Figure 1 Density ranges (bars) and averages (diamonds) for common Earth materials. The materials are grouped into categories as shown at the bottom of the chart. Within each category, average densities increase to the right. Data are from **Table 3**.

during flight. Low-pass filtering is required to remove nonlinear instrument drift along the line. Conventional tie-line leveling (Mittal, 1984), which minimizes misties at line crossings, and microleveling (Minty, 1991), which removes residual along-line noise, are also applied.

Both systems rely on an existing digital terrain model (DTM) extending 10 km beyond the survey area for use in terrain correction of the gravity tensor data. The Falcon system uses laser scanners to map the terrain in a swath beneath the aircraft and thereby improve the detail in the existing DTM. The uncorrected (free-air) gravity gradient data will be highly correlated with topography (**Figure 2(a)** and **2(b)**).

A variety of approaches, including along-line integration, vertical integration in the Fourier domain, and equivalent source methods, can be used to estimate vertical gravity g_z from the tensor components. Due to the unknown constant of integration, the estimated g_z may have an unknown base level, and regional wavelengths will also be missing. If gravimeter data are available, either from aircraft or from ground stations, a process known as 'conforming' (Dransfield, 2010) or 'crossover filtering' (Lane, 2004) can be used to combine short wavelengths from the calculated g_z with long wavelengths from the gravimeter data to improve the g_z estimate (**Figure 2(c)**).

11.12.8.4 Satellites

Seasat, in orbit during 1978, was the first satellite to include a radar altimeter for mapping sea-surface topography. The resulting improvements to the global marine geoid allowed computation of its first vertical derivative, the marine free-air gravity field, at wavelengths as short as 50 km and as long as 1000 km (Haxby et al., 1983). Sandwell and Smith (1997), using data from GEOSAT and ERS-1, increased the resolution down to 20 km wavelengths. **Figure 3** presents an example of satellite free-air marine gravity.

More recent satellite gravity missions (CHAMP, GRACE, and GOCE) have been designed to measure the global gravity field, over land and sea, but at longer wavelengths than the altimeter missions. These missions have already contributed to an improved global gravity model, and they will also result in improvements to the geoid (Bouman et al., 2013). The GOCE mission, which included an on-board gravity gradiometer, was designed to measure 1 mGal anomalies down to 100 km wavelengths. The GRACE mission, which involves tracking two chasing satellites, can measure wavelengths down to 600 km. This is adequate to monitor monthly groundwater storage changes in large aquifers (e.g., Strassberg et al., 2009).

11.12.9 Field Procedures on Land

Relative gravimeters will experience drift and thermal effects. They must remain on battery at all times before and during a survey in order to stabilize their internal temperature. To correct for drift, a local base station should be established and used as the first and last measurement station each day. The difference between the two daily base station measurements, after correcting for tidal effects, is considered to be due to linear instrument drift. The effects of Earth tides (**Chapter 3.06**) can be calculated (Longman, 1959) or measured at the base

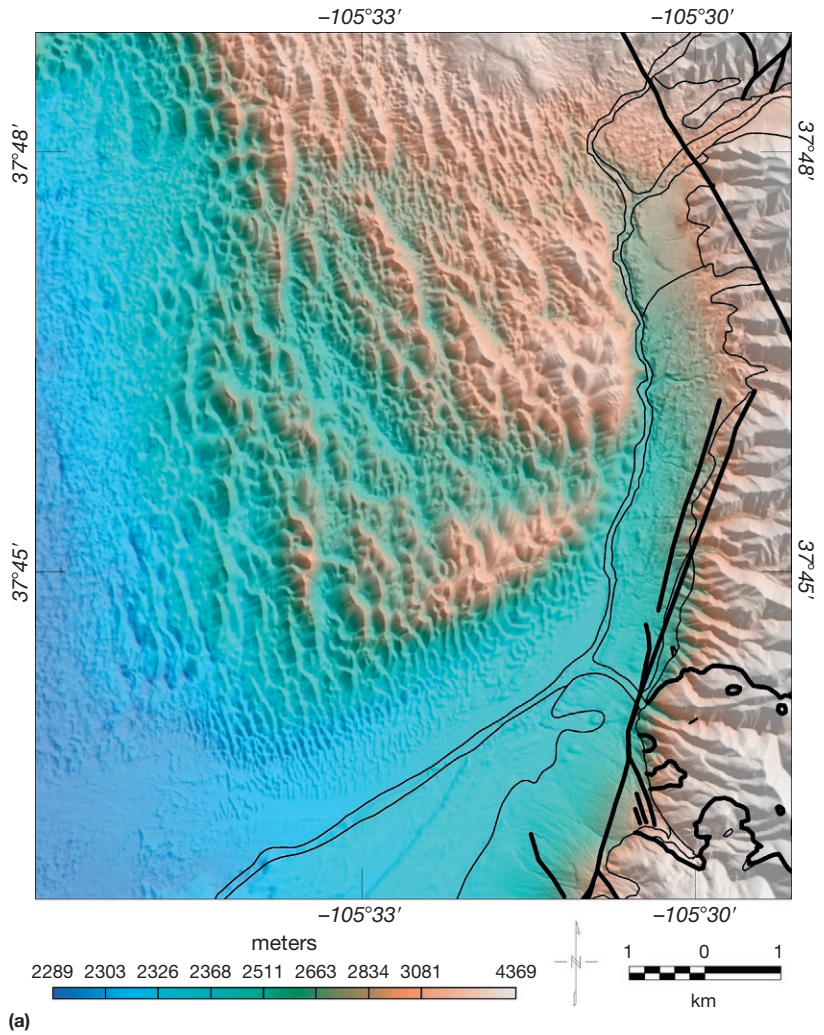


Figure 2 (Continued)

station. Thus, additional visits to the base station may be required during the day for ultraprecise measurements.

Local base stations should be tied to reference base stations at locations where the gravity is known. These locations are generally established by government agencies. A tie between a local base station and a reference base station will require at least three measurements, two at one base station and an intermediate reading at the other base station, in order to account for instrument drift. Additional back-and-forth readings can be used to check the accuracy of the local base station gravity value.

Gravity measurements should be taken at locations where the elevation is known to centimeter accuracy. These can be established benchmarks, spot elevations from topographic maps, or surveyed locations such as seismic shot-point locations. Modern gravity surveys often use differential GPS systems to obtain near-centimeter accuracy after postprocessing is applied. The latitude and longitude of the measurement location must also be known. These can be determined to sufficient accuracy from topographic maps or GPS measurements.

It is likely that gravity measurements already exist at most established benchmarks. Before planning a new gravity survey, it is best to check existing gravity databases to avoid repeating measurements unnecessarily.

Relative gravimeters must be leveled at each measurement location. The gravimeter is usually placed on a dish-shaped base plate with short tripod legs. Ground stability can be an issue here, especially on damp or frozen ground. Newer meters offer automatic leveling. Older meters require manual leveling using bubble levels and thumb screws on the legs.

After leveling, there is a waiting period while the damping mechanism stabilizes the spring. Readings are taken automatically or manually using an eyepiece focused on a lighted dial. Manual readings give values that are proportional to the relative vertical gravity; the proportionality constant is called the 'meter constant' and is different for each gravity meter. Automated readings usually correct for the meter constant along with other factors such as Earth tides and earthquake noise. It is a bad idea to drop or bump a relative gravity meter, as this can result in a "tare" in the meter readings, causing a step offset in the relative gravity values.

11.12.10 Data Reduction

Field measurements will establish the values of vertical gravitational acceleration at specific locations on the Earth. This represents a significant accomplishment, but further processing is required before these measurements can tell us anything

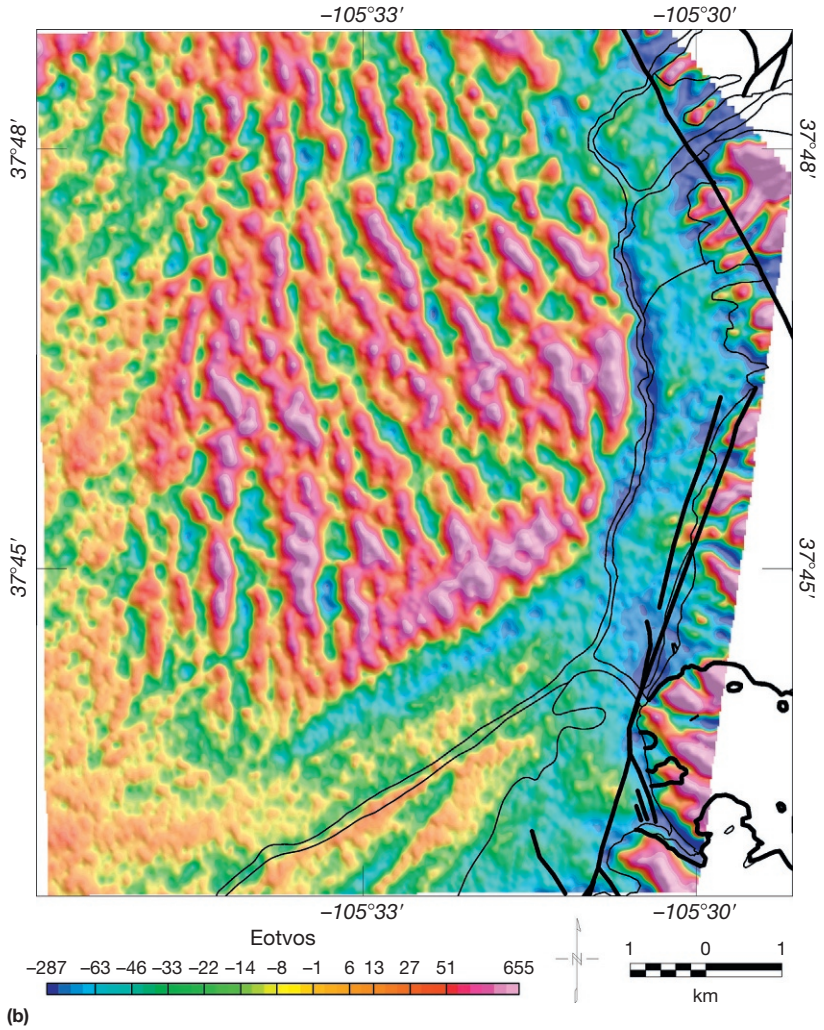


Figure 2 (Continued)

about the local density variations that are of interest to the exploration geophysicist.

To convert a gravity measurement at some location to a gravity anomaly value at the same location, it is necessary to subtract a term involving the theoretical gravity plus various correction terms.

The theoretical gravity on the reference ellipsoid is calculated from eqn [20] or [21]. It represents the latitude-dependent effect of rotation and distance from the center of mass of the Earth. At mid-latitudes, this latitude effect amounts to about 0.8138 mGal per kilometer, so a north–south mislocation of 1 m will produce a 0.008 mGal error.

11.12.10.1 Free-Air Anomalies

For gravity stations above sea level, the theoretical gravity must be reduced to account for the elevation difference between the station and sea level. If the station is at elevation h and the radius of the Earth at sea level is R, then the station is a distance $R+h$ from the center of mass. By the inverse square law, the difference between the theoretical gravity at sea level and the theoretical gravity at the station elevation should be

$$g_0 - g_h = g_0 - \frac{R^2}{(R+h)^2} g_0 = g_0 - \left(1 - 2\frac{h}{R} + 3\frac{h^2}{R^2}\cdots\right) g_0 \approx \frac{2hg_0}{R} \quad [25]$$

The factor $2g_0/R$ is 0.3067 mGal m^{-1} at the equator and 0.3093 at the pole. Generally, a value of 0.3086 mGal m^{-1}, representing the accepted 'standard gradient of gravity,' is used (LaFehr, 1991a).

The free-air gravity anomaly is defined as

$$g_{\text{faa}} = g_{\text{obs}} - (g_0 - 0.3086h) \quad [26]$$

where h is elevation in meters. For ocean-bottom or submarine measurements, h will be negative. The free-air anomaly represents the anomalous gravity at the station location that cannot be explained by the station latitude and the station elevation.

Free-air anomalies on land tend to be positive. This is due to the mass of the material underlying the station and extending above sea level. As a result of the large density

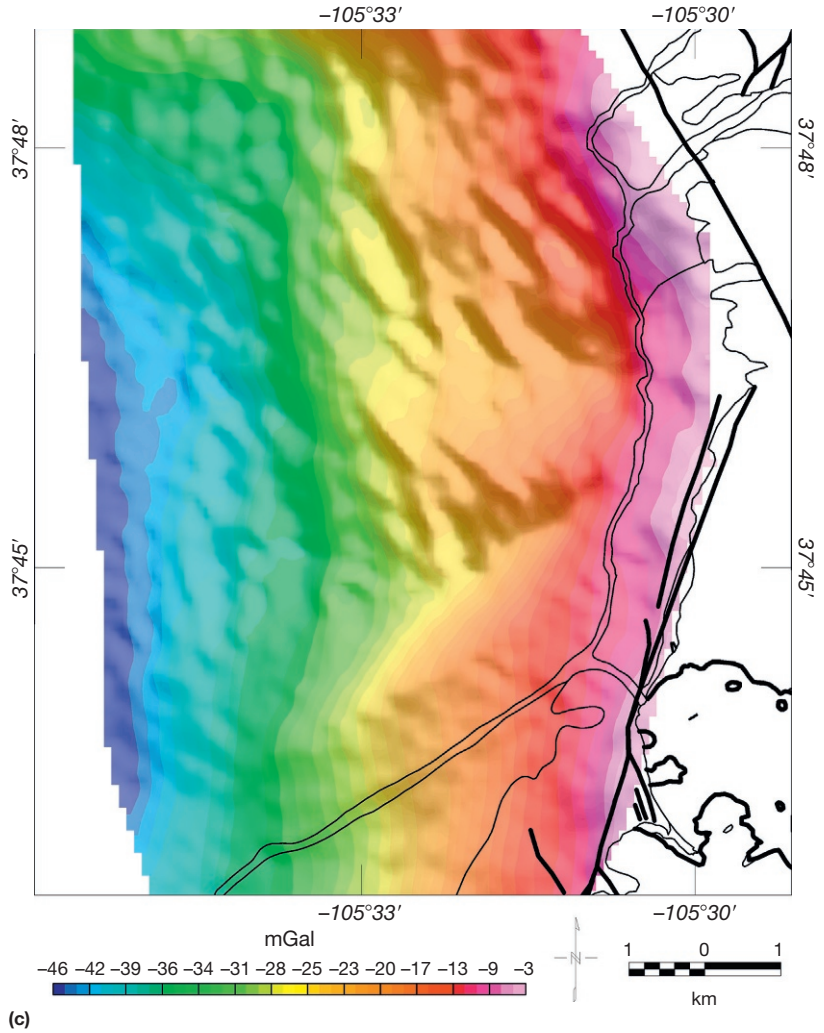


Figure 2 Results from a helicopter tensor gravity gradient survey of the Great Sand Dunes National Park, Colorado (Drenth et al., 2013): (a) the digital terrain model showing geologic contacts (thin lines) and faults (thick lines), (b) free-air vertical gravity gradient G_{zz}, and (c) conformed complete Bouguer anomaly g_z based on a terrain density of 2.67 g cm^{-3}. Note the short-wavelength content of G_{zz} compared to g_z. Courtesy of Ben Drenth, US Geological Survey.

contrast at the air–land interface and the proximity of this interface to the measurements, free-air anomalies also tend to be highly correlated with terrain. To properly model the free-air gravity anomaly, all density interfaces from the terrain or ocean surface down to the Moho and upper mantle need to be included.

11.12.10.2 Bouguer Anomalies

Exploration geophysicists are more interested in the gravity effect of anomalous subsurface mass than in the gravity effect of the terrain. Therefore, additional corrections are required to remove the terrain effects from the theoretical gravity. The mass lying between the station and the sea-level surface produces one part of the terrain effect. Stations at higher elevations will have more positive free-air gravity anomalies due to this mass. To the first order, the Bouguer slab formula [16] can be used to estimate and remove the effect of this mass. If we substitute the elevation of the station for the thickness of the slab, and use an average crustal density of 2.67 g cm^{-3} for the slab, then the simple Bouguer anomaly is defined as

$$g_{\text{sba}} = g_{\text{obs}} - (g_0 - 0.3086h + 0.11196h) \qquad [27]$$

For gravity data collected on surface ships at sea, the Bouguer correction is used to replace the ocean water by material of a higher density corresponding to that of the average sea floor. The intention is to remove the average gravity effect of the sea floor topography, so that all Bouguer anomalies will correspond to anomalous densities below the sea floor. In this case, the simple Bouguer anomaly is

$$g_{\text{sba}} = g_{\text{obs}} - (g_0 - 0.3086t + 0.0419(\rho - 1.03)(b + t)) \qquad [28]$$

where t is the tidal elevation above (or below) mean sea level, ρ is the chosen Bouguer density, and b is the bathymetry. The

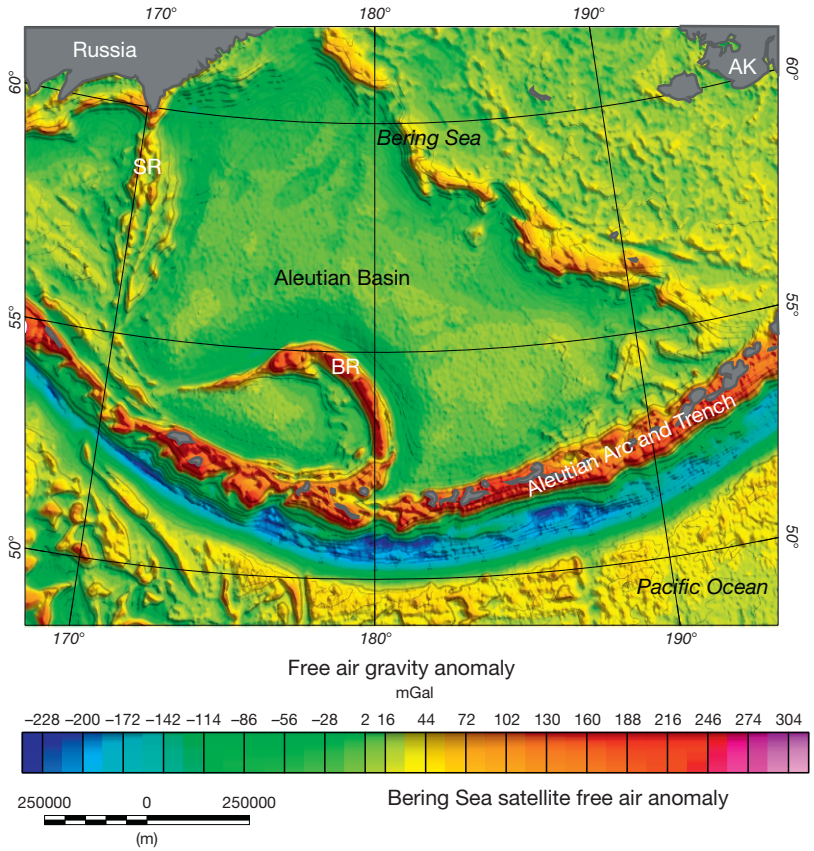


Figure 3 Marine free-air gravity anomaly (color image) and selected bathymetric contour lines from satellite altimetry (Sandwell and Smith, 2009) for the Aleutian Basin and surrounding region. The Aleutian volcanic arc and trench in the southern half of the figure form the tectonic plate boundary between the Pacific plate to the south and the North American plate to the north. The Aleutian basin is the deep water portion of the Bering Sea and includes the Bowers (BR) and Shirshov (SR) Ridges, both dated at about 30–33 My (late Oligocene). The tectonic history and development of this region of complex gravity and bathymetry are under active investigation for the assessment of potential hydrocarbon resources and for the establishment of sovereign territorial claims under Article 76 of the United Nations' Law of the Sea treaty. Courtesy of Rick Saltus, US Geological Survey.

tidal elevation is small and generally unknown in the open ocean, but it can be large, up to 12 m, in the near-shore (LaFehr and Nabighian, 2012).

More complete removal of terrain effects requires a 'complete Bouguer' correction consisting of three parts, the simple Bouguer correction that removes the effect of a Bouguer slab, a curvature correction that converts the slab into a spherical cap, and a terrain correction that removes the effect of the surrounding terrain.

11.12.10.2.1 Curvature correction

The simple Bouguer correction does not account for the curvature of the Earth. As a result, it removes nonexistent mass at elevations above and off to the sides of the station and it fails to remove mass just above sea level and off to the sides of the sea-level projection of the station. A correction for Earth curvature, called the Bullard B correction, converts the Bouguer slab correction to a correction for a spherical cap of thickness h and radius 166.735 km. The radius corresponds to the outer boundary of a standard zone used for isostatic compensation calculations (Hayford–Bowie zone O), and it was chosen by Bullard (1936) to minimize the average difference between the Bouguer slab anomaly and spherical cap anomaly over a large range of elevations. Following the work of Lambert (1930), LaFehr (1991b) derived the curvature correction in closed form and also provided the following approximation:

$$BB \approx 1.46308 \times 10^{-3} h - 3.52725 \times 10^{-7} h^2 + 5.1 \times 10^{-14} h^3 \quad [29]$$

The value of the Bullard B correction is 1.111 mGal at 1000 m elevation, and it achieves a maximum value of 1.518 mGal at 2100 m elevation. At elevations above 4150 m, the correction becomes negative.

Historically, the curvature correction has been ignored in the reduction of gravity data used by exploration geophysicists. Today, however, it has been incorporated into the Bouguer anomaly values appearing in standard public databases (Hildenbrand et al., 2002).

11.12.10.2.2 Terrain correction

Terrain surrounding a gravity station will have an effect on the gravity measurement. Adjacent topography that is higher than the station will counteract the downward pull of gravity, reducing the observed gravity value. Missing mass due to adjacent topography that is lower than the station will also reduce the downward pull of gravity and thus the observed gravity value.

The gravity effect of the surrounding terrain must be calculated from a model and added to the observed gravity in order to arrive at the gravity value that would be produced by flat terrain. In the past, this calculation was done using graphical overlays on topographic maps to estimate average elevations within rings about the station. Today, digital computers and digital elevation models are used to accomplish the same thing.

In the past, neither topographic maps nor available digital elevation models had the resolution necessary to calculate the terrain correction due to local topography very close to the station. For precise surveys, this 'inner zone' terrain correction had to be estimated in the field based on an examination of the local relief and knowledge of the gravity effect of topography based on model studies. Today, outside of frontier areas such as parts of Alaska, the available 10 and 3 m digital elevation models (DEMs) are adequate to calculate these innermost terrain corrections.

The average density of the terrain must be known for accurate terrain corrections. This value can be estimated from measured densities of rock samples, or it can be estimated directly from the free-air gravity anomalies in areas of topographic relief. A traditional approach called 'Nettleton profiling' involves collecting a gravity profile over a topographic feature and reducing the data using a range of densities; the density that produces a result that is least correlated with the terrain is considered to be a best estimate (Nettleton, 1939). An alternative approach that can be applied to both profile and scattered data involves modeling the free-air gravity as the sum of a term produced by uniform-density terrain and a polynomial term produced by deeper density contrasts. The optimum terrain density (and the polynomial coefficients) can be estimated by a least squares fit to the free-air anomaly. The number of polynomial coefficients can be varied until the correlation between the terrain and the free-air anomaly minus the polynomial term is maximized (see **Section 11.12.16.3**).

A simple Bouguer gravity anomaly value that has been corrected for curvature and terrain effects is called a 'complete' Bouguer anomaly. To model a complete Bouguer anomaly (CBA), only the anomalous masses underlying the observations need to be considered. The often large effect of the average-density topographic surface has been removed from the equation. For this reason, CBA values are preferred over free-air values for studies of anomalous near-surface densities in localized areas.

11.12.10.3 Isostatic Residual Anomalies

Complete Bouguer gravity anomalies in elevated land areas tend to be strongly negative. This is because isostatic compensation requires that elevated topography be compensated at depth by thicker crust. Crustal thickening at depth produces long-wavelength negative gravity anomalies because the crust is of lower density than the displaced mantle. Over the deep oceans, CBA tend to be strongly positive due to thinner crust and shallower mantle.

Isostatic corrections are used to remove the long-wavelength gravity anomalies produced by crustal thickening under the continents and crustal thinning under the oceans. The approach involves modeling the base of the crust as a density interface with a single density contrast, assigning a fixed crustal thickness at sea level, and balancing the mass in each vertical column down to a fixed compensation level somewhere below the deepest crust. If the sea-level crustal thickness is t_s, crustal density below sea level is ρ_c, and mantle density is ρ_m, then at a sea level site, the mass down to a compensation depth d_c is $t_s\rho_c + (d_c - t_s)\rho_m$. At a different site on land with elevation h, crustal density above sea level of ρ_t, and crustal thickness $h + t_h$, the mass is $h\rho_t + t_h\rho_c + (d_c - t_h)\rho_m$. Equating the two masses yields (Simpson et al., 1986)

$$t_h = t_s + \frac{h\rho_t}{\rho_m - \rho_c} \qquad [30]$$

A similar argument yields the elevation of the base of the crust under the ocean basins

$$t_h = t_s - b\frac{\rho_c - \rho_w}{\rho_m - \rho_c} \qquad [31]$$

where b is the sea water depth and ρ_w is the density of sea water. The gravity effect of the density contrast at the surface t_h is calculated using 3-D forward modeling and subtracted from the CBA values to generate the isostatic residual anomaly. Simpson et al. (1986) used values of $t_s = 30.0$ km, $\rho_t = \rho_c = 2.67$ g cm^{-3}, and $\rho_m - \rho_c = 0.35$ g cm^{-3} for their study of the conterminous United States.

11.12.10.4 Eötvös Correction

Gravity data collected on moving platforms, such as ships or aircraft, must be corrected for the motion of the platform. Because movement affects the apparent angular velocity of the Earth's rotation, gravitational attraction is reduced for platforms moving east and increased for platforms moving west. The effect is largest at the equator. The value of this Eötvös correction is (Glicken, 1962)

$$g_E = \left(\frac{R_\lambda + h}{R_\lambda^2}\right)\left(2R_\lambda \omega V \cos\lambda \sin\alpha + V^2\right) \qquad [32]$$

where λ is the geocentric latitude, R_λ is the radius of the Earth at that latitude, h is the height above the reference ellipsoid, ω is the angular velocity of rotation of the Earth, V is the velocity of the platform, and α is the heading (clockwise from north). If we ignore h and use the mean radius of the Earth, then for velocity in knots and g_E in mGal,

$$g_E = 7.5027V \cos\lambda \sin\alpha + 0.004154V^2 \qquad [33]$$

For a ship at the equator moving at 10 knots to the east, the Eötvös effect will cause a 75.44 mGal reduction in the observed gravity. The higher-order term becomes increasingly important at aircraft speeds.

11.12.11 Gravity Profiles and Grids

A gravity anomaly value taken at a single location provides little useful information unless it is combined with anomaly values taken at surrounding locations. The interpretation of gravity data generally requires multiple anomaly values organized into profiles or grids.

Although it is possible to interpret nonequispaced gravity profile data using forward modeling, many profile filtering and interpretation options require equispaced data. Generally, a nonequispaced profile is converted to an equispaced profile using spline interpolation techniques (Bhattacharyya, 1969).

For similar reasons, gravity data collected at scattered locations need to be organized into a regular grid prior to interpretation. A minimum-curvature gridding algorithm (Briggs, 1974) is generally used to interpolate scattered gravity data to a regular grid.

The original spacing of gravity stations along a profile or over an area will affect the interpretation of the data. The shortest wavelengths that can be reliably interpreted correspond to twice the average station spacing.

11.12.12 Filtering

Filtering of gravity data is a step that lies in between basic data reduction and advanced data interpretation. Filtering is used to remove noise, transform gravity field components, or isolate gravity anomalies of interest. Often, the goal of filtering is to produce a derivative product or map image that can be more easily interpreted than the observed gravity field. Filtering can be performed on profiles or on gridded gravity data, and it can be performed in the spatial domain or in the frequency domain through the use of Fourier transforms.

Spatial domain filtering can include low-pass or median filtering using convolution operators, polynomial surface fitting for regional-residual separation using linear least squares, smoothing using the diffusion equation, the calculation of horizontal gradients using first differences, or the calculation of curvature using local quadratic surface fitting.

Frequency (or Fourier) domain filtering can include upward or downward continuation from one level surface to another, calculation of horizontal or vertical derivatives, calculation of vertical integrals, transformations between gravity field components or between gravity field and gravity tensor components, and calculation of pseudo-magnetic fields from gravity fields and pseudo-gravity fields from magnetic fields. Matched filtering is a frequency domain band-pass filtering operation that attempts to separate an observed gravity field into components produced by horizontal density layers at multiple depths (Phillips, 2001; Syberg, 1972).

The pseudo-magnetic and pseudo-gravity transformations (Baranov, 1957; Blakely, 1995), based on Poisson's relation, serve to highlight the close relationship between behaviors of gravity and magnetic fields. This relationship allows magnetic interpretation methods to be applied to transformed gravity data and gravity interpretation methods to be applied to transformed magnetic data.

11.12.13 Interpretation: Forward Modeling

The calculation of gravity fields from density models is called forward modeling. To use this approach for interpretation, the calculated field must be compared to an observed gravity field, leading to iterative modification of the model in order to improve the fit. The simple gravity sources described earlier can be useful for forward modeling, but usually sources representing more complex geology are required.

11.12.13.1 2.5-D Polygonal Bodies

One of the easiest ways to interpret gravity profile data is through construction of forward models consisting of two-dimensional bodies striking perpendicular to the profile (along y) and having polygonal cross sections in the (x,z)-plane (**Figure 4(a)**). Such models can be built and modified rapidly on a computer. The vertical gravity field at the origin produced by such a body of density ρ defined by a closed N-sided polygon is given by (Blakely, 1995; Talwani et al., 1959)

$$g_z(0,0)=2\gamma\rho\sum_{n=1}^{N}\frac{\beta_n}{1+\alpha_n^2}\left[\log\frac{\sqrt{x_{n+1}^2+z_{n+1}^2}}{\sqrt{x_n^2+z_n^2}}-\alpha_n\left(\arctan\frac{z_{n+1}}{x_{n+1}}-\arctan\frac{z_n}{x_n}\right)\right] \quad [34]$$

Here, (x_n, z_n) represent the body corners in clockwise order and $x=\alpha_n z+\beta_n$ is the equation of a line passing through the nth side of the polygon. More precisely,

$$\alpha_n=\frac{x_{n+1}-x_n}{z_{n+1}-z_n},\quad \beta_n=x_n-\alpha_n z_n \quad [35]$$

Additional corrections can be applied to make the body strike limited or 2.5-D (Cady, 1980; Rasmussen and Pedersen, 1979).

11.12.13.2 Laminar Bodies

Three-dimensional bodies can be constructed from stacks of thin horizontal lamina (Talwani and Ewing, 1960) or from horizontal layers of uniform thickness (Plouff, 1975, 1976). In map view, the shape of each layer is defined by a polygon (**Figure 4(b)**). Calculations are similar to the 2-D case, with summations taken around the corners of each polygon. This approach was popular when computing resources were limited.

11.12.13.3 Gridded Surfaces

Many geologic interfaces can be represented by gridded surfaces. Examples include the topographic surface, the sea floor, sedimentary layer boundaries, the basement interface, and the Moho. Density layers bounded by gridded surfaces can be modeled by arrays of vertical rectangular prisms (**Figure 4(c)**). Each prism will have a constant density, but the prism densities can be made to vary horizontally throughout the layer. Gravity fields can be calculated by summing up the contributions of all the prisms. The vertical gravity of a rectangular prism extending from x_1 to x_2, y_1 to y_2, and z_1 to z_2, calculated at the origin, is given by (Blakely, 1995; Plouff, 1976)

$$g_z(0,0,0)=\gamma\rho\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\mu_{ijk}\left[z_k\arctan\frac{x_i y_j}{z_k R_{ijk}}-x_i\log\left(R_{ijk}+y_j\right)-y_j\log\left(R_{ijk}+x_i\right)\right] \quad [36]$$

where

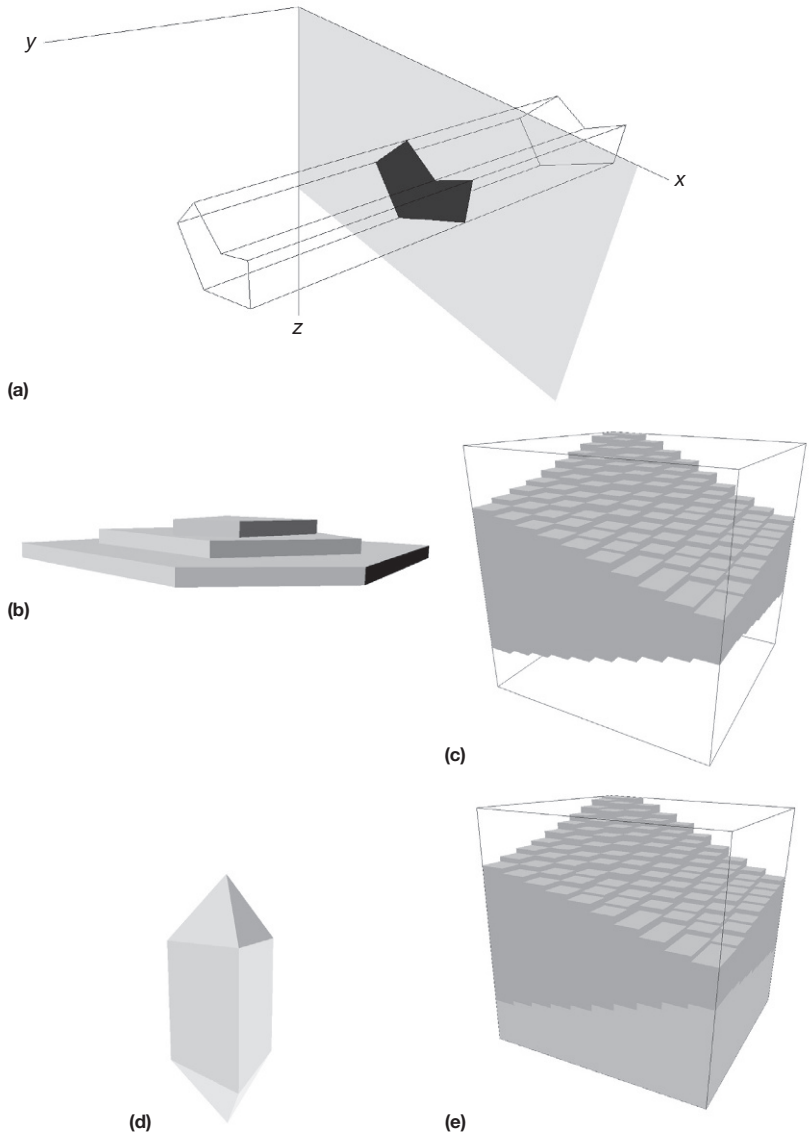


Figure 4 Some useful geometries for forward modeling. (a) 2-D or 2.5-D source having polygonal cross section and intersecting the plane below a gravity profile. (b) Horizontal lamina used to model topography. (c) A layer bounded by two surfaces and approximated using vertical rectangular prisms. (d) A 3-D body constructed out of polygonal facets. (e) A two-layer voxel model.

$$\mu_{ijk} = (-1)^i(-1)^j(-1)^k, R_{ijk} = \sqrt{x_i^2 + y_j^2 + z_k^2} \qquad [37]$$

The regular nature of gridded surfaces also makes them useful for Fourier transform operations. Parker (1972) showed that the gravity field on a horizontal surface at elevation z_0 produced by a layer with uneven top and bottom surfaces $z_1(x,y)$ and $z_2(x,y)$, containing horizontally varying density $\rho(x,y)$, could be rapidly calculated as

$$\Im[g_z(x,y,z_0)] = 2\pi\gamma e^{-|k|z_0}\sum_{n=1}^{\infty}\frac{|k|^{n-1}}{n!}\Im\left[\rho(x,y)\left(z_1^n(x,y) - z_2^n(x,y)\right)\right] \qquad [38]$$

where $\Im$ represents the two-dimensional Fourier transform operator and $|k| = \sqrt{k_x^2 + k_y^2}$ is the radial wavenumber. The main restriction of this approach is that the calculated field must be on a horizontal surface above all the sources, which makes it useful for sea-surface measurements, but not for measurements on topography. Parker (1995, 1996) extended his method to allow for a more general observation surface.

11.12.13.4 Faceted Bodies

Complex three-dimensional bodies or surfaces can be constructed out of polyhedral facets (**Figure 4(d)**). Details for calculating the gravity fields of these facets in the spatial domain can be found in Barnett (1976) and Okabe (1979). In the Fourier domain, the gravity field of a polyhedral facet can be expressed as a summation over the N vertices (x_n, y_n, z_n) of the polyhedron:

$$\Im[g_z] = -2\pi\gamma\rho\sum_{n=1}^{N}\alpha_n e^{-i\left(k_x x_n + k_y y_n\right) - |k| z_n} \quad [39]$$

Here, the α_n coefficients depend on the orientations of the two edges intersecting at vertex n (Blakely, 1995; Hansen and Wang, 1988). The Fourier transform of the gravity field of the complete body constructed out of the polyhedral facets has a form identical to [39] but with the α_n coefficients now depending on the orientations of all three edges intersecting at vertex n.

11.12.13.5 Voxel Models

Right rectangular prisms of uniform density can be combined to model complex density layers bounded by arbitrary surfaces (**Figure 4(e)**). Density can vary laterally within each layer. For calculations done in the space domain, the top and bottom of each prism can correspond to gridded surfaces, such as the topography, basement, or other geologic interfaces, and the field can be calculated on an arbitrary surface above the sources.

Unfortunately, space-domain calculations are slow. Fourier domain calculations, while fast, introduce several restrictions. If one- or two-dimensional Fourier transforms in the horizontal plane are used, the top and bottom of each layer must be a flat surface, and the calculated field must be on a flat surface above the sources. These restrictions preclude using this approach to construct accurate density models from ground-based gravity or draped airborne gravity in areas of significant topography.

Some of the above restrictions can be relaxed by using three-dimensional Fourier transforms in a volume. The 3-D Fourier transform of density in a volume $\hat{\rho}$ is related to the 3-D Fourier transform of the vertical gravity field in the same volume $\hat{g}_z$ by

$$\hat{g}_z\left(k_x, k_y, k_z\right) = i4\pi\gamma\frac{k_z}{|\mathbf{k}|^2}\hat{\rho}\left(k_x, k_y, k_z\right), |\mathbf{k}|^2 = k_x^2 + k_y^2 + k_z^2 \quad [40]$$

Here, $\mathbf{k}$ is the wavenumber vector (Caratori Tontini et al., 2009). This approach avoids the need for a flat observation surface, because the field volume can be sampled on any arbitrary surface. It also partially removes restrictions on surfaces bounding the layers, although the modeled surfaces will necessarily contain steps conforming to the vertical sample interval used for the volume.

11.12.14 Interpretation: Source Location Methods Applied to Gravity

Methods for locating isolated sources of potential fields require accurate calculations of the first and sometimes second derivatives of the field. For this reason, these methods have mainly been developed for and applied to aeromagnetic data, either along flightlines or in gridded form. The scattered nature of traditional gravity station data on land precludes the accurate calculation of derivatives.

Gravity profiles collected by surface ships at sea are suitable for source location analysis, as are airborne gravity data. Airborne gravity gradiometry provides direct measurements of the required derivatives.

The examples here cover two popular methods for quantitative interpretation of profile data. In general, these methods assume a simple two-dimensional source type and try to locate isolated sources of that type beneath the gravity profile. The strike of the source is assumed to be perpendicular to the direction of the profile. Both methods have extensions to three-dimensional sources and non-isolated (multiple) sources.

These methods generally require derivatives and integrals of the gravity profile data. Horizontal derivatives can be computed in the space domain using differences, splines, or curve fitting. Vertical derivatives and integrals are computed in the wavenumber domain using Fourier transform methods. It is useful to know that, for profiles crossing 2-D sources, the horizontal and vertical derivatives represent a Hilbert transform pair. Therefore, the horizontal derivative can be computed in the space domain, and the vertical derivative can be computed from the horizontal derivative using a Hilbert transform. Analysis done in small data windows using both the profile and its Hilbert transform will generally lead to more stable results than analysis done on the profile alone.

11.12.14.1 Werner Deconvolution

One of the earliest profile methods is due to Werner (1953) and has become known as Werner deconvolution (Hartman et al., 1971). When applied to gravity profile data, this approach assumes that the vertical gravity anomaly $G(x)$ on the level surface $z=0$ can be represented by the anomaly of a generalized horizontal density line source at location (x_0, z_0) plus a polynomial interference term of order n:

$$G(x) = \frac{A(x-x_0) + Bz_0}{(x-x_0)^2 + z_0^2} + C_0 + C_1 x + \dot{c} + C_n x^n \quad [41]$$

Here, A, B, and C_j are unknown constants. This can be rewritten as a polynomial with $n+5$ unknown coefficients a_i and b_i:

$$Gx^2 = a_0 + a_1 x + a_2 x^2 + \dot{c} + a_n x^n + a_{n+1}x^{n+1} + a_{n+2}x^{n+2} + b_0 G + b_1 Gx \quad [42]$$

where

$$b_0 = \left(-x_0^2 - z_0^2\right),\ b_1 = 2x_0 \quad [43]$$

Therefore, by selecting $n+5$ consecutive points along the profile, and solving for the coefficients of the polynomial, one can estimate the location of the line source as

$$x_0 = \frac{1}{2}b_1,\ z_0 = \frac{1}{2}\sqrt{-4b_0 - b_1^2} \quad [44]$$

If the same analysis is done on the horizontal or vertical derivative of the gravity profile, then the source locations represent the tops or edges of density sheets instead of lines. If the analysis is done on the second horizontal or second vertical derivative of the gravity profile, then the source locations represent density contact sources. In practice, a linear or quadratic interference polynomial is assumed, so either six or seven points along the profile are used in each analysis. Furthermore, it is common practice to increase the window size by using every other point, every third point, and so on in order to search for deeper sources. Results are often simplified by

averaging solutions that occur close together and discarding isolated solutions.

The Hilbert transform $H(x)$ of $G(x)$ can be incorporated into the analysis by adding a second equation at each point of the data window:

$$Hx^2 = c_0 + c_1 x + c_2 x^2 + \dot{c} + c_n x^n + c_{n+1} x^{n+1} + c_{n+2} x^{n+2} + b_0 H + b_1 Hx \quad [45]$$

This allows us to reduce the window length to $n+3$ points and leads to a system of $2n+6$ equations with $2n+6$ unknowns.

$$\begin{bmatrix} G_1 & G_1x_1 & 1 & 0 & x_1 & 0 & \cdots & x_1^{n+2} & 0 \\ H_1 & H_1x_1 & 0 & 1 & 0 & x_1 & \cdots & 0 & x_1^{n+2} \\ G_2 & G_2x_2 & 1 & 0 & x_2 & 0 & \cdots & x_2^{n+2} & 0 \\ H_2 & H_2x_2 & 0 & 1 & 0 & x_2 & \cdots & 0 & x_2^{n+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ G_{n+3} & G_{n+3}x_{n+3} & 1 & 0 & x_{n+3} & 0 & \cdots & x_{n+3}^{n+2} & 0 \\ H_{n+3} & H_{n+3}x_{n+3} & 0 & 1 & 0 & x_{n+3} & \cdots & 0 & x_{n+3}^{n+2} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ a_0 \\ c_0 \\ a_1 \\ c_1 \\ \vdots \\ a_{n+2} \\ c_{n+2} \end{bmatrix}$$

$$= \begin{bmatrix} G_1x_1^2 \\ H_1x_1^2 \\ G_2x_2^2 \\ H_2x_2^2 \\ \vdots \\ G_{n+3}x_{n+3}^2 \\ H_{n+3}x_{n+3}^2 \end{bmatrix} \quad [46]$$

Under this approach, Werner deconvolution is done on the gravity profile and its Hilbert transform to find line solutions, on the horizontal derivative of the gravity profile and its Hilbert transform (the vertical derivative) to find sheet solutions, and on the second horizontal derivative of the gravity profile and its Hilbert transform (the second vertical derivative) to find contact sources. In the case of gravity gradiometer profiles, G_{zx} and G_{zz} are the measured horizontal and vertical derivatives of the vertical gravity field and can be used to find sheet solutions, provided G_{zy} is small.

With some additional effort (Hansen and Simmonds, 1993), Werner deconvolution can be extended to solve for multiple sources in the window, provided that they all are of the same type (lines, sheets, or contacts). There is even a three-dimensional version of the method based on isolated (Nabighian and Hansen, 2001) or multiple (Hansen, 2005) sources.

11.12.14.2 Euler Deconvolution

A second analysis method, called Euler deconvolution, is based on Euler's homogeneity relationship. A function $V(x, y, z)$ is homogeneous of degree n if it satisfies Euler's equation

$$(x-x_0)\frac{\partial V}{\partial x} + (y-y_0)\frac{\partial V}{\partial y} + (z-z_0)\frac{\partial V}{\partial z} = -nV \quad [47]$$

Furthermore, if a function is homogeneous of degree n, its spatial derivatives are homogeneous of degree $n+1$, and its spatial integrals are homogeneous of degree $n-1$. If we plug our equations for the vertical gravity of a point mass [5] and its derivatives into eqn [47], we find that the gravity field is homogeneous of degree 2. It follows that the gravity field of a density line source is homogeneous of degree 1, that of a density sheet source is homogeneous of degree 0, and that of a density contact is homogeneous of degree -1. Because the degree of homogeneity of a gravity field is related to the geometry of the source, it has been given a special name – the structural index.

Euler deconvolution (Reid et al., 1990; Thompson, 1982) is based on a reordering of Euler's equation to put the unknown source coordinates on the one side of the equation and the known observation coordinates and an assumed structural index on the other side of the equation and adding a term to account for an unknown constant in the field. Thus, if the observed field G is the true field of the source G' plus an unknown constant C', then the Euler equation becomes

$$(x-x_0)\frac{\partial G}{\partial x} + (y-y_0)\frac{\partial G}{\partial y} + (z-z_0)\frac{\partial G}{\partial z} = -nG' = -n(G-C') \quad [48]$$

Reordering gives

$$x_0\frac{\partial G}{\partial x} + y_0\frac{\partial G}{\partial y} + z_0\frac{\partial G}{\partial z} + C = x\frac{\partial G}{\partial x} + y\frac{\partial G}{\partial y} + z\frac{\partial G}{\partial z} + nG \quad [49]$$

where $C = nC'$. If we know the field and its derivatives at several locations, say in a small window along a profile or over a grid, then eqn [49] becomes a linear system of equations that can be solved for the unknown source location and constant C.

For a gravity profile, we assume that the sources are two-dimensional and strike along the y-axis. This simplifies eqns [48] and [49] by removing the y-derivative terms. If anomalies from shallow and deep sources interfere, removing a constant from the field may not provide adequate anomaly separation. We can borrow an idea from Werner deconvolution and incorporate higher-order interference terms. If the observed field G contains a linear interference term and is measured on the level $z = 0$, then our system of equations for Euler deconvolution of profile data becomes

$$x_0\frac{\partial G}{\partial x} + z_0\frac{\partial G}{\partial z} + C_0 + C_1 x = x\frac{\partial G}{\partial x} + nG \quad [50]$$

The four unknowns on the left can be solved for using linear least squares within data windows of various lengths along the profile. A good least-squares routine should provide an estimate of the standard error for each parameter, so solutions with large percentage errors in depth can be rejected. It is also a good idea to reject solutions with x-coordinates that fall outside the data window being analyzed.

If we choose to analyze both the gravity profile $G(x)$ and its Hilbert transform $H(x)$, then [50] is supplemented by a second equation for each point in the m-point data window:

$$x_0\frac{\partial H}{\partial x} + z_0\frac{\partial H}{\partial z} + D_0 + D_1 x = x\frac{\partial H}{\partial x} + nH \quad [51]$$

This leads to a system of 2 m equations with six unknowns:

$$\begin{bmatrix} \frac{\partial G_1}{\partial x} & \frac{\partial G_1}{\partial z} & 1 & 0 & x_1 & 0 \\ \frac{\partial H_1}{\partial x} & \frac{\partial H_1}{\partial z} & 0 & 1 & 0 & x_1 \\ \frac{\partial G_2}{\partial x} & \frac{\partial G_2}{\partial z} & 1 & 0 & x_2 & 0 \\ \frac{\partial H_2}{\partial x} & \frac{\partial H_2}{\partial z} & 0 & 1 & 0 & x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial G_m}{\partial x} & \frac{\partial G_m}{\partial z} & 1 & 0 & x_m & 0 \\ \frac{\partial H_m}{\partial x} & \frac{\partial H_m}{\partial z} & 0 & 1 & 0 & x_m \end{bmatrix} \begin{bmatrix} x_0 \\ z_0 \\ C_0 \\ D_0 \\ C_1 \\ D_1 \end{bmatrix} = \begin{bmatrix} x_1 \frac{\partial G_1}{\partial x} + nG_1 \\ x_1 \frac{\partial H_1}{\partial x} + nH_1 \\ x_2 \frac{\partial G_2}{\partial x} + nG_2 \\ x_2 \frac{\partial H_2}{\partial x} + nH_2 \\ \vdots \\ x_m \frac{\partial G_m}{\partial x} + nG_m \\ x_m \frac{\partial H_m}{\partial x} + nH_m \end{bmatrix} \quad [52]$$

The three-dimensional version of Euler deconvolution is particularly suited to application to gravity gradient tensor data, because all required derivatives of the gravity field are available directly (Zhang et al., 2000). The multiple-source version of Euler deconvolution (Hansen and Suciu, 2002) requires higher-order derivatives of the field; therefore, it appears to be less practical than multiple-source Werner deconvolution, which requires only the first derivatives.

11.12.15 Interpretation: Inversion

11.12.15.1 Trial and Error

An efficient forward modeling routine can form the basis for inversions based on trial-and-error fitting. The 2.5-D polygonal cross-sectional approach to profile modeling has proven especially useful in this regard. A user interface that allows rapid repositioning of polygonal corners and rapid editing of both corners and body densities coupled with a feedback display of the anomaly misfit is helpful for such analysis (Saltus and Blakely, 1983, 1993; Webring, 1985).

11.12.15.2 Iterative Inversions for Density Interfaces

11.12.15.2.1 Basin thickness from gravity profile data: Bott's method

One of the first practical uses of gravity inversion was for estimating the thickness of sedimentary basins. It is an application that continues to be useful today. Bott (1960) developed an iterative method for estimating basin thickness from gravity profile data. The basin fill is assumed to have a constant density contrast $\Delta\rho$ with the underlying basement, and the sources are approximated by 2-D right rectangular prisms centered under each gravity measurement. The basin thickness t_j under each gravity station is initialized using the Bouguer slab formula

$$t_j^{(0)} = \frac{g_j}{2\pi\gamma\Delta\rho}, j = 1, 2, \ldots, N \quad [53]$$

At each iteration k, the gravity field of the prism model $g_j^{(k)}$ is calculated at the station locations and subtracted from the observed gravity anomaly to get a residual $g_j - g_j^{(k)}$. The thickness is updated:

$$t_j^{(k+1)} = t_j^{(k)} + \frac{\left(g_j - g_j^{(k)}\right)}{2\pi\gamma\Delta\rho} \quad [54]$$

And the iterations continue until the residual is small.

11.12.15.2.2 Density interface from gravity profile data: The Parker–Oldenburg method

As an alternative to Bott's method, Oldenburg (1974) rearranged a one-dimensional version of the Fourier transform-based forward eqn [38] of Parker (1972) to produce the inverse expression

$$\Im[h(x)] = -\frac{1}{2\pi\gamma\rho}\Im[g_z(x)]e^{|k|z_0} - \sum_{n=2}^{\infty}\frac{|k|^{n-1}}{n!}\Im[h^n(x)] \quad [55]$$

Here, the solution $h(x)$ represents the bottom surface of a density layer with a flat top surface at $z=0$, and the measurements $g_z(x)$ are assumed to be on a flat surface $z=z_0$ above the layer. The expression is applied iteratively to update $h(x)$ from an initial guess, such as a constant. Low-pass filtering may be required to ensure convergence. In theory, the approach could easily be extended to the 2-D case.

11.12.15.2.3 Density interfaces from gridded gravity: The Cordell–Henderson method

Using a slightly different approach, Cordell and Henderson (1968) extended Bott's method to estimate the shape of a 3-D density interface from a gridded gravity anomaly. The sources were vertical prisms of uniform density centered under each grid node, and they could extend down from a horizontal reference surface for a basin as in Bott's method, up from a horizontal surface at fixed depth for a dome or seamount, or both up and down an equal amount from a horizontal surface at fixed depth for an ore body. If the prism thicknesses were initialized as in Bott's method, then the ratio of the observed field to the calculated field was used to update the thickness

$$t_j^{(k+1)} = \frac{g_j}{g_j^{(k)}} t_j^{(k)} \quad [56]$$

The iterations continue until the thickness of the prisms is stable.

11.12.15.2.4 Basin/basement gravity separation and basin thickness: The Jachens–Moring method

The gravity field over a closed sedimentary basin consists of a component due to sedimentary fill in the basin and a component due to density variations in the underlying basement. The Jachens–Moring method (Jachens and Moring, 1990) provides an iterative approach to separate the two components and estimate the basin thickness. The original approach involves using the basin boundary to separate the gravity stations that are on the basin fill from the gravity stations that are on the surrounding basement rocks. An initial basement gravity field is generated by gridding only the gravity anomalies at stations located on basement and interpolating the field across the area of the basin. An initial basin gravity field is generated by subtracting the gridded basement gravity field from the gridded observed gravity field and assigning any positive values to zero. Cordell–Henderson inversion is used to convert the initial basin gravity field into an initial basin thickness model. The basin thickness model is used to calculate the revised basin gravity field at the locations of the basement gravity stations, and this field is subtracted from the observed gravity values to improve the basement gravity estimates at these stations.

The modified basement gravity values are gridded as before, and the iterative process continues until changes to the basin thickness are small. A modified approach that uses only gridded gravity anomaly data and treats the grid nodes as gravity station locations also works well. The Cordell–Henderson inversion can be modified to accommodate a stepped density–depth relationship instead of a single density contrast. The density–depth relationship can vary with position in the basin to accommodate known basin heterogeneities. Seismic interpretations and boreholes can be used as constraints on the configuration of the basin. This requires modeling the gravity field of the basin at the seismic stations and borehole locations and subtracting the modeled gravity from the observed gravity to generate virtual basement gravity stations.

11.12.15.3 Physical Property Inversions

This section addresses the problem of estimating subsurface densities directly from gravity measurements or gravity gradient measurements on the ground or sea surface, in aircraft, or in boreholes. This problem is non-unique; there are many density distributions that will fit a given set of measurements. On the other hand, there are many density distributions that will not fit the measurements. The trick is to impose constraints that lead to a set of solutions that fit the data and are geologically reasonable (Li and Oldenburg, 1998; Oldenburg and Li, 2005).

All physical property inversions start from the forward problem, which in the case of gravity or gravity gradiometry is linear

$$\mathbf{g} = \mathbf{A}\boldsymbol{\rho} \qquad [57]$$

The gravity anomaly (and/or gradient) values in vector $\mathbf{g}$ are related to the densities in vector $\boldsymbol{\rho}$ by the coefficients in matrix $\mathbf{A}$. In most cases, we do not actually calculate the field using matrix algebra. Instead, we use one of the forward modeling approaches mentioned earlier.

Because there are usually many more unknown densities in the model than there are field measurements, linear inverse theory (Menke, 1989) suggests that we use the 'underdetermined' or 'data-space' solution

$$\boldsymbol{\rho} = \boldsymbol{\rho}_0 + \mathbf{P}\mathbf{A}^{\mathrm{T}}\left(\mathbf{A}\mathbf{P}\mathbf{A}^{\mathrm{T}} + \mathbf{D}\right)^{-1}\left(\mathbf{g} - \mathbf{A}\boldsymbol{\rho}_0\right) \qquad [58]$$

Here, I have introduced a starting density model $\boldsymbol{\rho}_0$, a model weighting matrix $\mathbf{P}$, and a data covariance matrix $\mathbf{D}$. Conventional matrix notation is used: $\mathbf{A}^{\mathrm{T}}$ is the transpose of matrix $\mathbf{A}$, and $\mathbf{A}^{-1}$ represents the matrix inverse. Generally, $\mathbf{P}$ and $\mathbf{D}$ are assumed to be diagonal matrices, meaning they are filled with zeros except along the diagonal. The diagonal elements of $\mathbf{D}$ are normally constant, unless the data quality or data type is variable. The real work is done by the diagonal elements of $\mathbf{P}$, which can be used to impose constraints such as depth weighting and compactness (Last and Kubik, 1983; Portniaguine and Zhdanov, 1999). Another important constraint for gravity inversion is the requirement that the resulting densities be positive. One way to handle this is through a change of variable from $\boldsymbol{\rho}$ to $\sqrt{\boldsymbol{\rho}}$ or to $\log \boldsymbol{\rho}$. The density range can be further limited using bound constraints such as the logarithmic barrier method (Oldenburg and Li, 2005).

Practical inversion algorithms require a nonlinear approach to the solution of [58] using the conjugate gradient or similar method to solve for the inverse matrix at each iteration, a line search to ensure convergence, and possible update of matrix $\mathbf{P}$ to implement compactness. Normally, all large matrix operations can be replaced by forward calculations [57]. These can be done in the space domain or using fast Fourier transforms for improved speed. Parallel processing is beneficial in order to fully explore the solution space in a reasonable amount of time.

11.12.16 Field Examples

11.12.16.1 Estimating Overburden Thickness and Bedrock Depth at the Tooele Army Depot, Utah, Using Constrained Gravity Inversion

The configuration of the bedrock under the Tooele Army Depot (TEAD), Tooele, Utah, helps to contain one or more contaminant plumes of saline groundwater. These plumes are considered to be environmental hazards by the US Environmental Protection Agency. The knowledge of the bedrock surface comes from boreholes and gravity inversion. The U.S. Geological Survey (USGS) was asked to estimate the thickness of the overburden and the elevation of the bedrock surface using gravity inversion while, at the same time, forcing the result to agree with the available borehole data. The boreholes are concentrated over a bedrock high in the center of the study area, but there are scattered boreholes throughout the study area, many of which do not reach the bedrock surface. The bedrock also outcrops along the southern boundary of the study area.

A critical early step in the inversion involved the removal of the regional gravity field, represented as a planar surface, from the CBA data. The parameters of this regional field were estimated using the known depth to bedrock in 49 boreholes and at four outcrop locations. A density of 2.20 g cm^{-3} was assumed for the overburden and a density of 2.67 g cm^{-3} was assumed for the bedrock. The known depth to bedrock was translated into an estimated overburden gravity anomaly at the 53 locations using the Bouguer slab approximation with a density contrast of -0.47 g cm^{-3}. The estimated overburden gravity was subtracted from the observed CBA to give estimates of the regional gravity field at the 53 locations. A planar surface was then fit to these 53 estimates of regional gravity using a linear least-squares approach. This planar surface was assumed to represent the regional gravity field, and it was subtracted from the gridded CBA data to produce the observed residual gravity anomaly grid.

The gravity inversion was done using a modified Cordell–Henderson approach (Cordell and Henderson, 1968). In this modified approach, the initial Bouguer slab thickness estimate for iteration zero and the thickness estimates for subsequent iterations calculated using eqn [56] were modified to agree with the borehole data before each new gravity field was calculated. The modification involved a correction surface constructed from data at the borehole locations as follows:

- The preliminary overburden thickness grid produced by eqn [56] was sampled at the locations of the boreholes and outcrops. Differences between the true bedrock depths (where known) or minimum bedrock depths (as represented by the total depths of boreholes that did not intersect bedrock) and the sampled values were computed. For boreholes (or outcrops) that intersected the bedrock surface, the difference was retained regardless of sign. For boreholes that did not intersect the bedrock surface, the difference was retained if it was negative (the preliminary surface was above the bottom of the borehole) and set to zero if it was positive (the preliminary surface was below the bottom of the borehole). The edited difference was then gridded using minimum curvature to produce the correction surface.

This correction surface was added to the preliminary overburden thickness grid to produce the revised overburden thickness grid at each iteration.

Following iteration one, additional iterations did not produce any real improvement to the result (Table 4), so the overburden thickness estimate from iteration one was added to the topographic surface grid generated from the 2 arcsecond DEM to produce the final estimated bedrock surface elevation (Figure 5).

The final thickness estimate included undulations north of the mountain front in the southeast corner of the area. These undulations are largely unconstrained by any gravity or borehole data. The gravity field calculated from the final thickness estimate was in good agreement with the observed residual gravity, having an RMS misfit of 0.59 mGal (Table 4). This surface is likely to be unreliable in the southwest and northwest corners and elsewhere where borehole data are absent and the gravity stations are sparse. Based on these results, the U.S. Army Corps of Engineers made significant modifications to their groundwater model and later performed verification drilling.

11.12.16.2 Estimating Thickness of Sedimentary Rocks in the Yukon Flats Basin, East-Central Alaska, Using Constrained Iterative Gravity Inversion

The Yukon Flats basin of east-central Alaska is a potential source of hydrocarbon resources for local consumption and possible export. The knowledge of the subsurface configuration of the basin is largely restricted to a few proprietary seismic reflection profiles covering a limited area. The seismic profiles were reprocessed and reinterpreted by the U.S. Geological Survey in preparation for an assessment of the oil and gas resources of the basin (Stanley et al., 2004). The assessment effort required knowledge of the basin configuration between and away from the seismic profiles.

Table 4 Results of modified Cordell–Henderson gravity inversion at the Tooele Army Depot

Iteration	*Misfit in overburden thickness at bedrock boreholes and outcrops*		*Gravity RMS misfit (mGal)*
	Mean (ft)	*Standard deviation (ft)*	
0	−5.4	12.6	0.84
1	−6.6	15.3	0.59
2	−5.1	13.9	0.61
3	−5.0	15.3	0.60
4	−5.4	16.1	0.60

To extend the interpretation of the basin thickness across the entire area of the basin, a version of iterative Jachens–Moring gravity inversion (Jachens and Moring, 1990) based on gridded gravity data was used. At each iteration, the gravity field was separated into basin (Figure 6) and basement (Figure 7) components, and the basin component was inverted to estimate thickness. The inversion was easily constrained to agree with the interpreted basement surface along the seismic profiles. A single density contrast of $-0.17\ \mathrm{g\,cm^{-3}}$, based on a 2-D model along one seismic profile, was used in the inversion. The effects of lowering the density contrast to $-0.13\ \mathrm{g\,cm^{-3}}$, a median value for 2-D models along all seismic profiles, was explored but rejected. The smaller density contrast increased basin thickness away from the seismic profiles but produced unstable results in the Crooked Creek sub-basin.

Sedimentary depocenters deep enough to act as 'kitchens' for the generation of hydrocarbons were a primary target of the investigation. In addition to the main depocenter interpreted from the seismic data with over 8 km of fill, the preferred gravity inversion indicates four apparent depocenters with over 4 km of fill in the western Yukon Flats basin, a similar depocenter in the eastern Yukon Flats basin, and a large depocenter with over 7 km of fill in the Crooked Creek sub-basin (Figure 8). However, unexposed felsic intrusions in the basement could be responsible for some of the gravity lows interpreted as depocenters; granitic rocks are exposed near the western end of the basin. Results for the Crooked Creek sub-basin are consistent with magnetic and magnetotelluric modeling, but this and the other apparent depocenters require further testing by drilling and seismic profiling.

11.12.16.3 Estimating Terrain Density from Gravity Profile Data, National Petroleum Reserve Alaska

The knowledge of the average density of the terrain is required for optimum terrain correction during Bouguer reduction. It is also useful for initializing densities during gravity inversion and for estimating near-surface elastic constants from measured seismic refraction velocities.

Figure 9 shows a free-air gravity profile collected at shot points along a seismic line in the National Petroleum Reserve Alaska (NPRA). The free-air gravity is strongly correlated with the terrain except at the north end of the line. To estimate the average density of the terrain along the line, we assume that the free-air gravity anomaly consists of two parts: (1) a largely short-wavelength part represented by a constant X times the terrain, $z(x)$, and (2) a longer-wavelength part produced by deeper density contrasts and represented by a polynomial.

$$g_{faa}(x_i) = Xz(x_i) + A + \sum_{j=1}^{M} B_j x_i^j;\ i = 1, N \qquad [59]$$

The first part is completely correlated with terrain; the second part is uncorrelated with terrain.

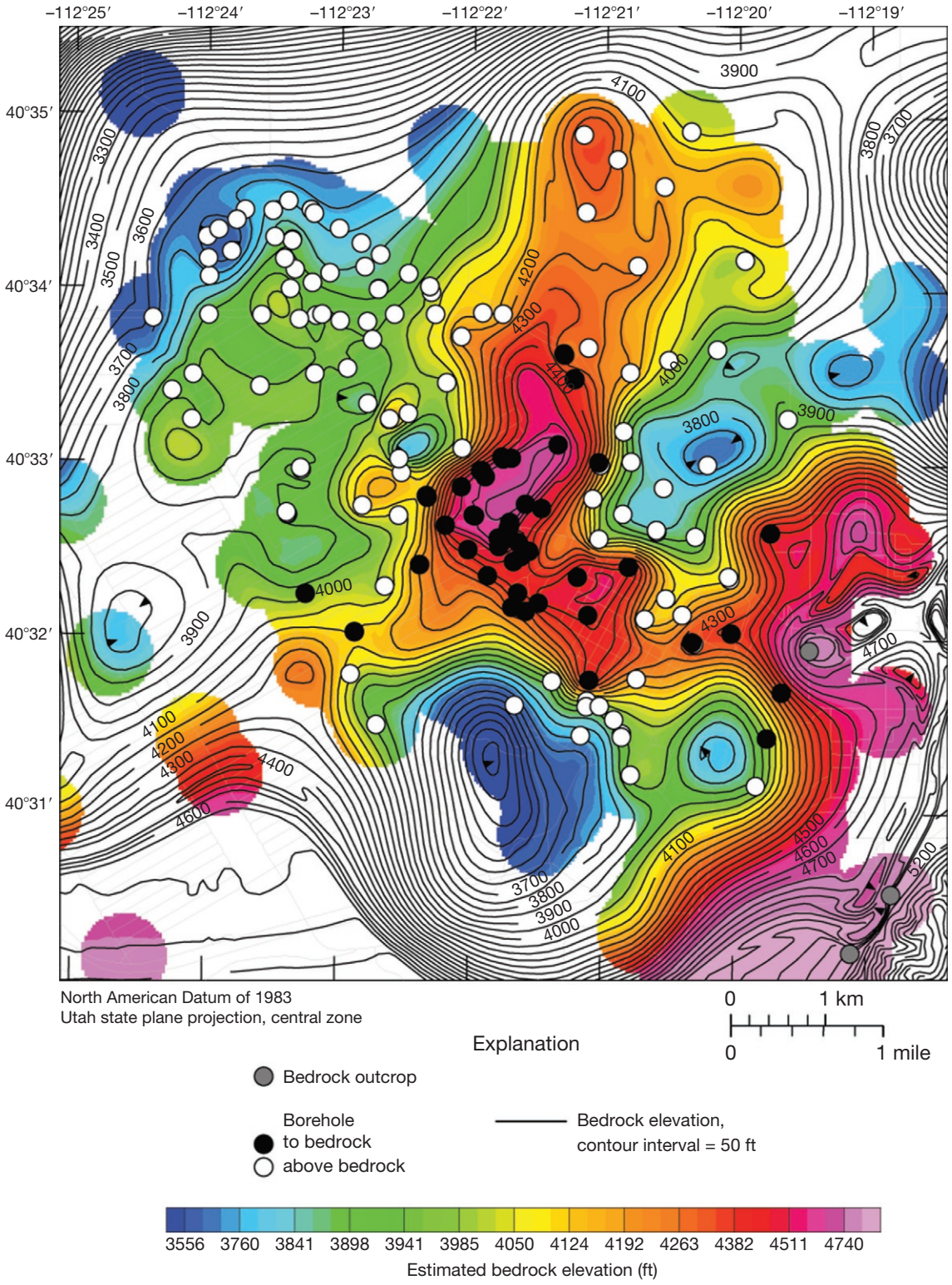


Figure 5 Bedrock elevation for the Tooele Army Depot, Tooele, Utah, as estimated using constrained Cordell–Henderson gravity inversion. The solution is colored within 1500 ft of outcrops, boreholes, and gravity stations. Note the nonlinear color scale. Reliability is expected to be poor in the uncolored areas.

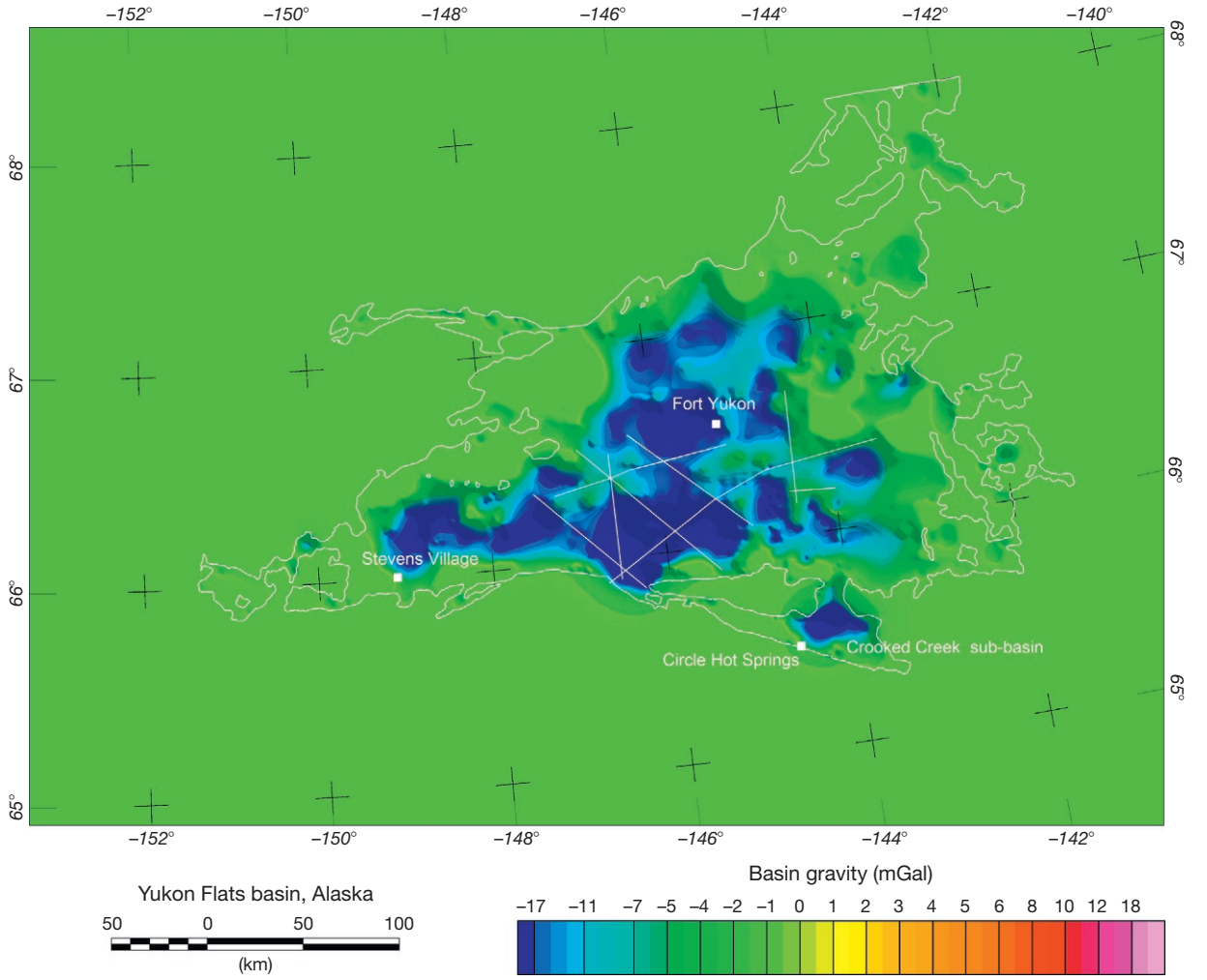


Figure 6 The vertical gravity component due to basin fill in the Yukon Flats basin, east-central Alaska, as estimated using modified Jachens–Moring gravity inversion. The basin boundary and the locations of seismic reflection profiles used to constrain the solution are shown as white lines. Local settlements are shown as labeled white squares. The gravity field is displayed in color shaded relief, and the nonlinear color scale is taken from **Figure 7**.

We can use linear least squares to estimate X and the polynomial coefficients A, B_j for polynomials of various degrees. The optimum polynomial degree M is chosen to maximize the correlation coefficient between the terrain gravity $g_T(x)$ and the terrain $z(x)$:

$$\mathrm{corr}(g_T, z) = \frac{\sum_{i=1}^{N}(g_T(x_i) - \overline{g}_T)(z(x_i) - \overline{z})}{\sqrt{\sum_{i=1}^{N}(g_T(x_i) - \overline{g}_T)^2}\sqrt{\sum_{i=1}^{N}(z(x_i) - \overline{z})^2}} \quad [60]$$

where

$$g_T(x_i) = g_{faa}(x_i) - A - \sum_{j=1}^{M} B_j x_i^j \quad [61]$$

Coefficient X can be converted into a terrain density ρ_B using the Bouguer slab formula

$$X = 2\pi\gamma\rho_B \quad [62]$$

The approach also yields an approximate CBA

$$\hat{g}_{cba}(x_i) = g_{faa}(x_i) - Xz(x_i) \quad [63]$$

consisting of the polynomial term plus any signal resulting from variations in the terrain density.

For the data of **Figure 9**, an average terrain density of 2.37 g cm^{-3} results from using a 12° polynomial fit.

11.12.16.4 Using Time-Lapse Gravity to Monitor Subsurface Water Storage in a Former Coal Mine, Arvada, Colorado

The Leyden mine in Arvada, Colorado, was an active underground coal mine until 1958, when it was closed and converted into a natural gas storage facility. After 40 years, gas storage was suspended, and the remaining gas was removed through water displacement. The mine and the water injection well were then converted into a water storage facility by the City of Arvada.

Students and faculty at the Colorado School of Mines (Davis et al., 2008) first used model studies to show that the density increase caused by water injected into the mine could be detected using microgravity surveys. They then decided to

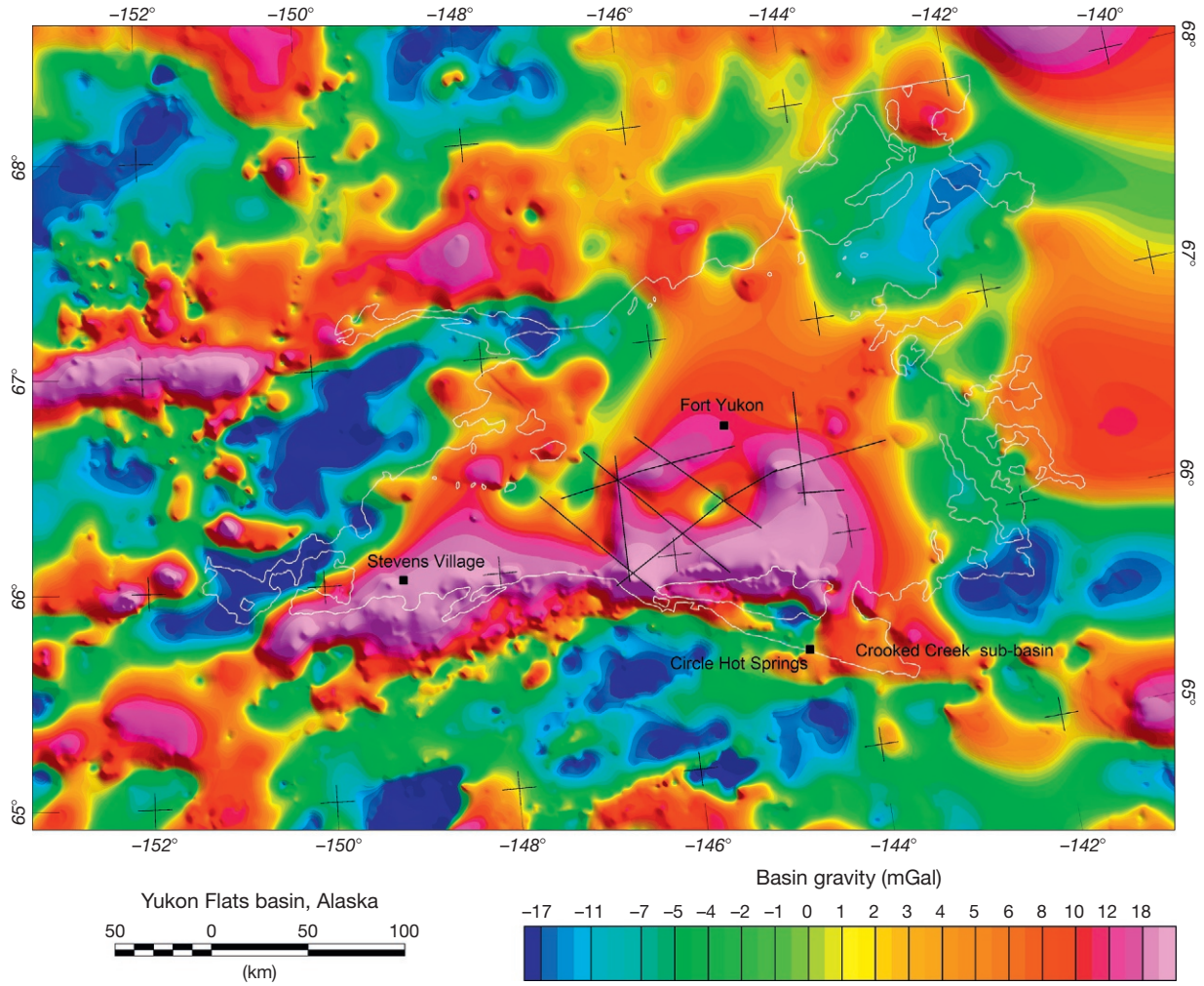


Figure 7 The vertical gravity component due to basement sources in the vicinity of the Yukon Flats basin, east-central Alaska, as estimated using modified Jachens–Moring gravity inversion. The basin boundary is shown as a white line. The locations of seismic reflection profiles are shown as black lines. Local settlements are shown as labeled black squares. The gravity field, which represents the complete Bouguer anomaly with a 200 km high-pass filter applied, is displayed in color shaded relief with a nonlinear color scale.

conduct a time-lapse gravity survey in the southern part of the mine area, which contained the injection well. The goals of the survey were to map the distribution and movement of water through the mine, identify areas of leakage and areas of ceiling collapse, and recommend locations for recovery wells.

The mine workings occupy about 4 km^2 within a 16 km^2 area at a depth of 300 m. Microgravity measurements were made at 56 gravity stations in April 2004, October 2004, and February 2005. The gravity stations over the mine were spaced 300 m apart, corresponding to the depth to the mine. Three of the gravity stations were well away from the mine area and were used as reference stations. During the survey, time-lapse measurements at the reference stations varied by less than 5 μGal.

The largest increase in gravity between April and February (Figure 10) occurred at stations just northwest of the injection well (195 μGal) and at stations north of the injection well on the northern edge of the survey (155 μGal). These stations are directly over mined areas, and the increased gravity is attributed to water filling of areas enlarged through roof collapse. Gravity did not increase significantly outside of the mined area, indicating that leakage was not a problem.

11.12.17 Summary

The gravitational method is based on Newton's law of gravitation. Instruments used to measure gravity include relative gravimeters, absolute gravity meters, gradiometers, and satellites. The vertical component of the gravity field $g_z(\mathbf{r})$ is the quantity measured by most gravimeters and used in most interpretations of subsurface density variations. Gravity gradients, which are normally presented as tensor components, can be measured by newer airborne instruments. These offer higher resolution at the shorter wavelengths of interest for near-surface investigations.

The physical property associated with the gravitational method is density or density contrast. Common near-surface Earth materials have densities that range from 1.0 to near 4.0 g cm^{-3}. Sample densities of rocks can be measured in the

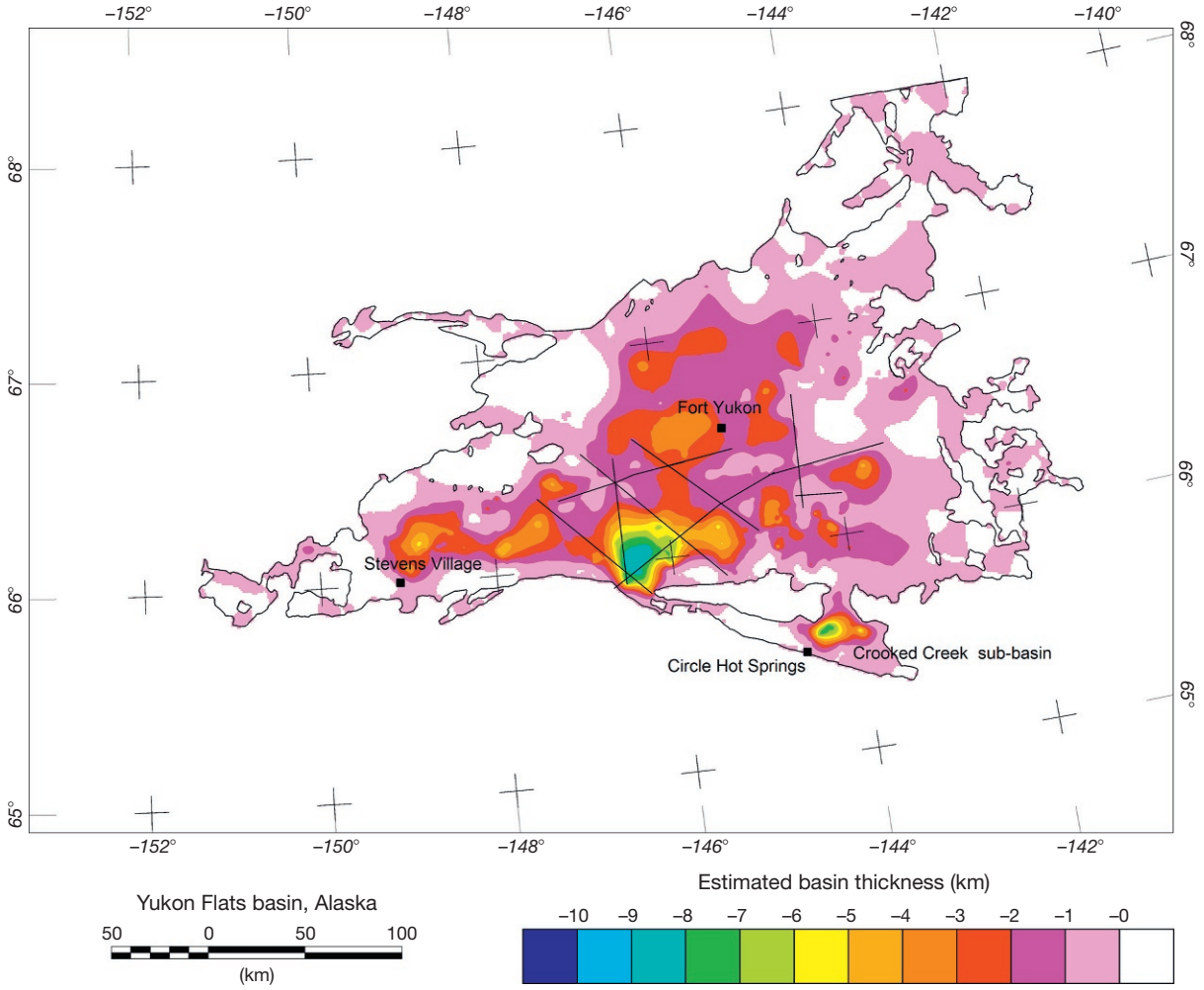


Figure 8 Thickness of sedimentary fill in the Yukon Flats basin, east-central Alaska, as estimated using modified Jachens–Moring gravity inversion. The basin boundary and the locations of seismic reflection profiles are shown as black lines. Local settlements are shown as labeled black squares. White areas indicate zero thickness.

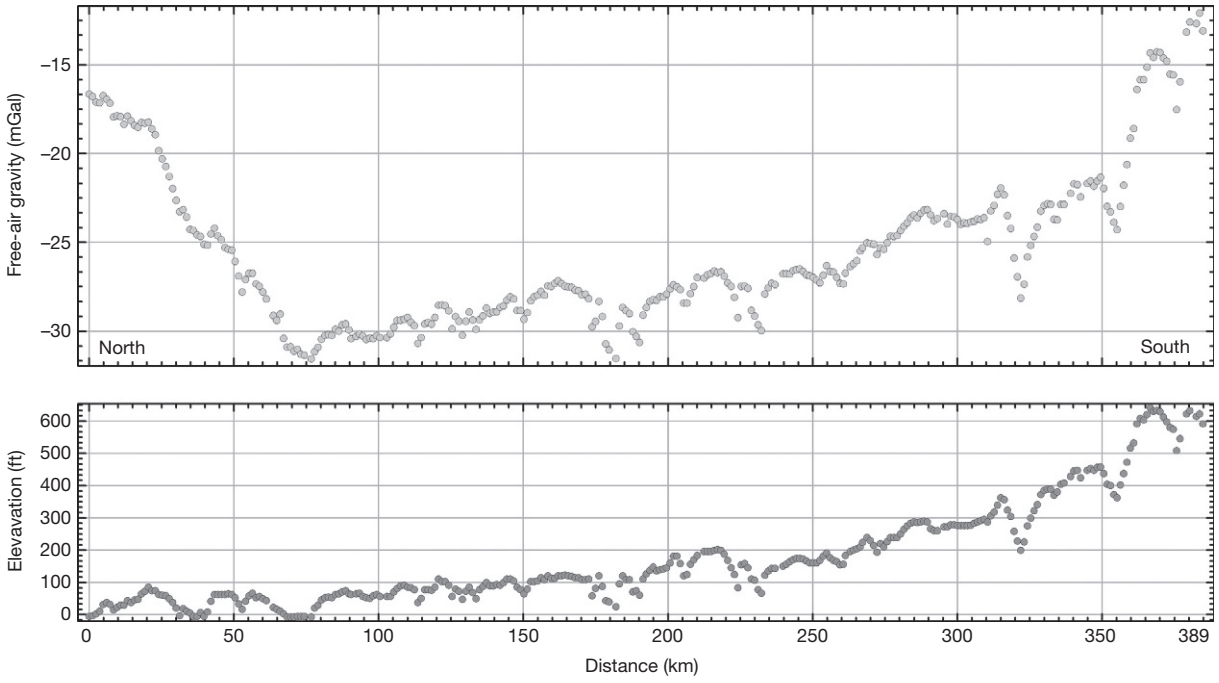


Figure 9 Free-air gravity (top) and elevation (bottom) along NPRA seismic line 127-76. The high correlation between the curves allows the average density of the terrain to be estimated. Gravity data from Saltus RW, et al. (2008).

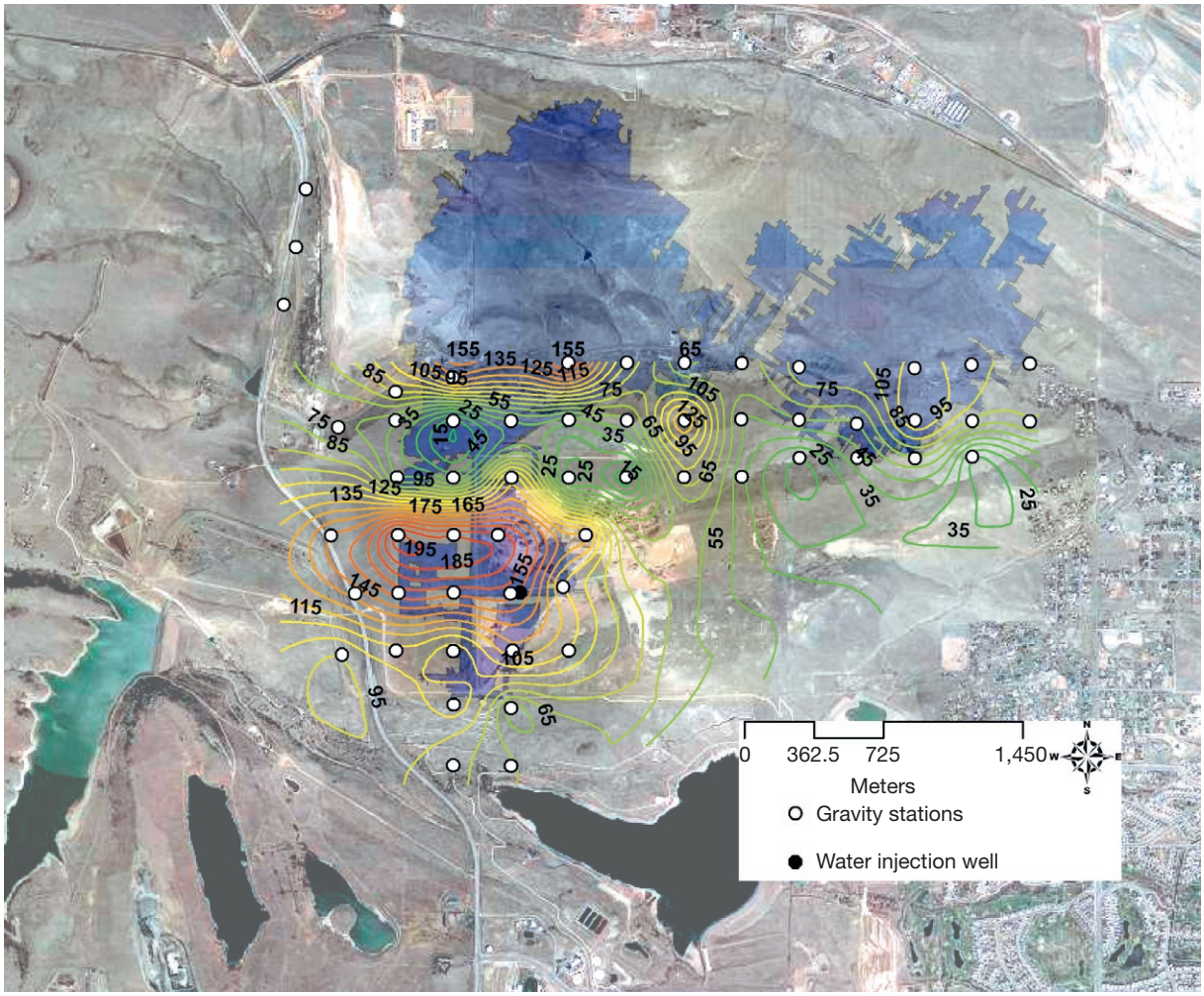


Figure 10 Results of a time-lapse microgravity survey over the Leyden mine underground water storage area in Arvada, Colorado (Davis et al., 2008). Contours indicate the change in gravity in μGal between April 2004 and February 2005 due to water injection. Reoccupied gravity stations are shown by white circles. The water injection well is indicated by a white triangle. The blue pattern shows the extent of the underground mine at a depth of 300 m. The figure, courtesy of Kristofer Davis, University of British Columbia, has been previously published in Geophysics and is copyright by the Society of Exploration Geophysicists.

laboratory and bulk densities can be measured in a borehole using a gamma–gamma logging tool or a borehole gravimeter. Densities of surface materials can be estimated from correlations between topography and free-air gravity.

To convert a gravity measurement at some location to a gravity anomaly value at the same location, it is necessary to subtract terms involving the theoretical gravity plus various correction terms. The free-air anomaly represents the anomalous gravity at the station location that cannot be explained by the station latitude and the station elevation. Additional corrections remove the terrain effects and produce Bouguer anomalies. These reflect only the gravity effect of anomalous subsurface mass. Isostatic corrections are used to remove the long-wavelength gravity anomalies produced by crustal thickening under the continents and crustal thinning under the oceans.

The interpretation of gravity data generally requires multiple anomaly values organized into profiles or grids. Filtering can be performed on profiles or on gridded gravity data to remove noise, transform gravity field components, or isolate gravity anomalies of interest.

To use forward modeling for interpretation, the calculated field of the density model must be compared to an observed gravity field, leading to iterative modification of the model in order to improve the fit. Forward models consisting of two-dimensional bodies having polygonal cross sections can be used to interpret gravity profile data. Three-dimensional bodies can be constructed from stacks of thin horizontal lamina, assembled from arrays of vertical rectangular prisms, constructed out of polyhedral facets, or represented by volume-filling voxels.

Two popular source location methods for quantitative interpretation of gravity profile data, with extensions to three-dimensional sources, are Werner deconvolution and Euler deconvolution. Other interpretation methods include iterative inversions for density interfaces, illustrated here by field examples, and physical property inversions for variable distributions of density. The future of the gravitational method will probably

be defined by improvements in airborne gravity gradiometry. As the noise levels of these instruments are reduced through technological improvements, their numbers are increased above the current handful, and the cost of gradiometer surveys goes down, we should see gravity gradient surveys supplementing high-resolution aeromagnetic surveys as a cost-effective near-surface mapping tool.

Disclaimer

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Baranov V (1957) A new method for interpretation of aeromagnetic maps: Pseudo-gravimetric anomalies. *Geophysics* 22: 359–383.

Barnett CT (1976) Theoretical modeling of the magnetic and gravitational fields of an arbitrarily shaped three-dimensional body. *Geophysics* 41: 1353–1364.

Bell RE and Hansen RO (1998) The rise and fall of early oil field technology: The torsion balance gradiometer. *The Leading Edge* 17: 81–83.

Bhattacharyya BK (1969) Bicubic spline interpolation as a method for treatment of potential field data. *Geophysics* 34: 402–423.

Blakely RJ (1995) *Potential Theory in Gravity & Magnetic Applications*. Cambridge: Cambridge University Press.

Bott MHP (1960) The use of rapid digital computing methods for direct gravity interpretation of sedimentary basins. *Geophysical Journal of the Royal Astronomical Society* 3: 63–67.

Bouman J, Floberghagen R, and Rummel R (2013) More than 50 years of progress in satellite gravimetry. *Eos, Transactions American Geophysical Union* 94: 262–270.

Briggs IC (1974) Machine contouring using minimum curvature. *Geophysics* 39: 39–48.

Bullard EC (1936) Gravity measurements in East Africa. *Philosophical Transactions. Royal Society of London* 235(757): 486–497.

Cady JW (1980) Calculation of gravity and magnetic anomalies of finite-length right polygonal prisms. *Geophysics* 45: 1507–1512.

Caratori Tontini F, Cocchi L, and Carmisciano C (2009) Rapid 3-D forward model of potential fields with application to the Palinuro Seamount magnetic anomaly (southern Tyrrhenian Sea, Italy). *Journal of Geophysical Research* 114: B02103. http://dx.doi.org/10.1029/2008JB005907.

Cordell L and Henderson RG (1968) Iterative three-dimensional solution of gravity anomaly data using a digital computer. *Geophysics* 33: 596–601.

Davis K, Li Y, and Batzle M (2008) Time-lapse gravity monitoring: A systematic 4D approach with application to aquifer storage and capacity. *Geophysics* 73: WA61–WA69.

Dobrin MB (1960) *Introduction to Geophysical Prospecting*. New York: McGraw-Hill Book Company.

Dranesfield MH and Lee JB (2004) The FALCON airborne gravity gradiometer system. In: Lane R (ed.) Airborne Gravity 2004—Abstracts from the ASEG-PESA Airborne Gravity Workshop. Geoscience Australia Record 2004/18, 15–19.

Dransfield MH (2010) Conforming Falcon gravity and the global gravity anomaly. *Geophysical Prospecting* 58: 469–483.

Drenth BJ, Abraham JD, Grauch VJS, and Hodges G (2013) Digital data from the Great Sand Dunes airborne gravity gradient survey, south-central Colorado. U.S. Geological Survey Open-File Report 2013-1011.

Glicken M (1962) Eötvös corrections for a moving gravity meter. *Geophysics* 27: 531–533.

Hansen RO (2005) 3D multiple-source Werner deconvolution for magnetic data. *Geophysics* 70: L45–L51.

Hansen RO and Simmonds M (1993) Multiple-source Werner deconvolution. *Geophysics* 58: 1792–1800.

Hansen RO and Suciu L (2002) Multiple-source Euler deconvolution. *Geophysics* 67: 525–535.

Hansen RO and Wang X (1988) Simplified frequency-domain expressions for potential fields of arbitrary three-dimensional bodies. *Geophysics* 53: 365–374.

Hartman RR, Teskey DJ, and Friedberg JL (1971) A system for rapid digital aeromagnetic interpretation. *Geophysics* 36: 891–918.

Haxby WF, Karner GD, LaBrecque JL, and Weissel JK (1983) Digital images of combined oceanic and continental data sets and their use in tectonic studies. *Eos, Transactions American Geophysical Union* 64: 995–1004.

Hildenbrand TG, Briesacher A, Flanagan G, et al. (2002) Rational and operational plan to upgrade the U.S. gravity database. U.S. Geological Survey Open-File Report 02-463.

Jachens RC and Moring BC (1990) Maps of thickness of Cenozoic deposits and the isostatic residual gravity over basement for Nevada. U.S. Geological Survey Open-File Report 90-404.

Kellogg OD (1953) *Foundations of Potential Theory*. New York: Dover.

LaFehr TR (1991a) Standardization in gravity reduction. *Geophysics* 56: 1170–1178.

LaFehr TR (1991b) An exact solution for the gravity curvature (Bullard B) correction. *Geophysics* 56: 1179–1184.

LaFehr TR and Nabighian M (2012) *Fundamentals of Gravity Exploration*. Tulsa: Society of Exploration Geophysicists.

Lambert WD (1930) The reduction of observed values of gravity to sea level. *Bulletin of Geodesy* 26: 107–181.

Lane R (2004) Integrating ground and airborne data into regional gravity compilations. In: Lane R (ed.) Airborne Gravity 2004 – Abstracts from the ASEG-PESA Airborne Gravity Workshop. Geoscience Australia Record 2004/18, 81–97.

Last BJ and Kubik K (1983) Compact gravity inversion. *Geophysics* 48: 713–721.

Lee JB (2001) FALCON gravity gradiometer technology. *Exploration Geophysics* 32: 247–250.

Li Y and Oldenburg DW (1998) 3-D inversion of gravity data. *Geophysics* 63: 109–119.

Longman IM (1959) Formulas for computing the tidal acceleration due to the moon and the sun. *Journal of Geophysical Research* 64: 2351–2355.

Menke W (1989) Geophysical Data Analysis: Discrete Inverse Theory (revised edition). *International Geophysics Series*, vol. 45. San Diego: Academic Press.

Minty BRS (1991) Simple micro-levelling for aeromagnetic data. *Exploration Geophysics* 22: 591–592.

Mittal PK (1984) Algorithm for error adjustment of potential field data along a network. *Geophysics* 49: 467–469.

Murphy CA (2004) The Air-FTG™ airborne gravity gradiometer system. In: Lane R (ed.) Airborne Gravity 2004—Abstracts from the ASEG-PESA Airborne Gravity Workshop. Geoscience Australia Record 2004/18, 7–14.

Nabighian MN, Ander ME, Grauch VJS, et al. (2005) The historical development of the gravity method in exploration. *Geophysics* 70: 63ND–89ND.

Nabighian MN and Hansen RO (2001) Unification of Euler and Werner deconvolution in three dimensions via the generalized Hilbert transform. *Geophysics* 66: 1805–1810.

Nettleton LL (1939) Determination of density for reduction of gravimeter observations. *Geophysics* 4: 176–183.

Newton I (1687) *Philosophie Naturalis Principia Mathematica*.

Okabe M (1979) Analytical expressions for gravity anomalies due to homogeneous polyhedral bodies and translations into magnetic anomalies. *Geophysics* 44: 730–741.

Oldenburg DW (1974) The inversion and interpretation of gravity anomalies. *Geophysics* 39: 526–536.

Oldenburg DW and Li Y (2005) Inversion for applied geophysics: A tutorial. In: Butler DK (ed.) *Near-Surface Geophysics*, pp. 89–150. Tulsa: Society of Exploration Geophysicists.

Parker RL (1972) The rapid calculation of potential anomalies. *Geophysical Journal of the Royal Astronomical Society* 31: 447–455.

Parker RL (1995) Improved Fourier terrain correction I. *Geophysics* 60: 1007–1017.

Parker RL (1996) Improved Fourier terrain correction II. *Geophysics* 61: 365–372.

Phillips JD (2001) Designing matched bandpass and azimuthal filters for the separation of potential-field anomalies by source region and source type. Australian Society of Exploration Geophysicists. In: *15th Geophysical Conference and Exhibition* Expanded Abstracts CD-ROM.

Plouff D (1975) Derivation of formulas and FORTRAN programs to compute gravity anomalies of prisms. National Technical Information Service No. PB-243-526, U.S. Department of Commerce.

Plouff D (1976) Gravity and magnetic fields of polygonal prisms and application to magnetic terrain corrections. *Geophysics* 41: 727–741.

Portniaguine O and Zhdanov MS (1999) Focusing geophysical inversion images. *Geophysics* 64: 874–887.

Rasmussen R and Pedersen LB (1979) End corrections in potential field modeling. *Geophysical Prospecting* 27: 749–760.

Reid AB, Allsop JM, Granser H, Millett AJ, and Somerton IW (1990) Magnetic interpretation in three-dimensions using Euler deconvolution. *Geophysics* 55: 80–91.

Saltus RW and Blakely RJ (1983) HYPERMAG, an interactive, two-dimensional gravity and magnetic modeling program. U.S. Geological Survey Open-File Report No. 83-241.

Saltus RW and Blakely RJ (1993) HYPERMAG, an interactive, 2- and 2 1/2-dimensional gravity and magnetic modeling program, version 3.5. U.S. Geological Survey Open-File Report No. 93-287.

Saltus RW, Brown PJ, Morin RL, and Hill PL (2008) 2006 compilation of Alaska gravity data and historical reports. U.S. Geological Survey Data Series DS264, CDROM.

Sandwell DT and Smith WHF (1997) Marine gravity anomaly from Geosat and ERS-1 satellite altimetry. *Journal of Geophysical Research* 102: 10039–10054.

Sandwell DT and Smith WHF (2009) Global marine gravity from retracked Geosat and ERS-1 altimetry: Ridge segmentation versus spreading rate. *Journal of Geophysical Research* 114: B01411. http://dx.doi.org/10.1029/2008JB006008.

Simpson RW, Jachens RC, Blakely RJ, and Saltus RW (1986) A new isostatic gravity map of the conterminous United States with a discussion on the significance of isostatic residual anomalies. *Journal of Geophysical Research* 91: 8348–8372.

Stanley RG, Ahlbrandt TS, Charpentier RR, et al. (2004) Oil and gas assessment of Yukon Flats, east-central Alaska, 2004. U.S. Geological Survey Fact Sheet 2004-3121.

Strassberg G, Scanlon BR, and Chambers D (2009) Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States. *Water Resources Research* 45: W05410. http://dx.doi.org/10.1029/2008WR006892.

Syberg FJR (1972) A Fourier method for the regional-residual problem of potential fields. *Geophysical Prospecting* 20: 47–75.

Talwani M and Ewing M (1960) Rapid computation of gravitational attraction of three-dimensional bodies of arbitrary shape. *Geophysics* 25: 203–225.

Talwani M, Worzel JL, and Landisman M (1959) Rapid gravity computations for two-dimensional bodies with application to the Mendocino submarine fracture zone. *Journal of Geophysical Research* 64: 49–59.

Thompson DT (1982) EULDPH: A new technique for making computer-assisted depth estimates form magnetic data. *Geophysics* 47: 31–37.

Webring M (1985) SAKI, A Fortran program for generalized linear inversion of gravity and magnetic profiles. U.S. Geological Survey Open-File Report No. 85-122.

Werner S (1953) Interpretation of magnetic anomalies at sheet-like bodies. *Sveriges Geologiska Undersok., Ser C C Arsbok*. 43: N06.

Zhang C, Mushayandebvu MF, Reid AB, Fairhead JD, and Odegard ME (2000) Euler deconvolution of gravity tensor gradient data. *Geophysics* 65: 512–520.

11.13 Tools and Techniques: Nuclear Magnetic Resonance

M Müller-Petke, Leibniz Institute for Applied Geophysics, Hannover, Germany
U Yaramanci, Leibniz Institute for Applied Geophysics, Hannover, Germany; Berlin Technical University, Berlin, Germany

© 2015 Elsevier B.V. All rights reserved.

11.13.1 Introduction 420
11.13.2 Fundamentals of NMR 421
11.13.2.1 Spin Magnetization 421
11.13.2.2 Excitation 421
11.13.2.3 Relaxation 422
11.13.3 Laboratory NMR 422
11.13.3.1 The Governing Concept 422
11.13.3.2 Free Induction Decay 423
11.13.3.3 The Transverse Relaxation Time T_2 423
11.13.3.4 The Longitudinal Relaxation Time T_1 425
11.13.3.5 Inverse Modeling: From Measured Data to Relaxation Times 425
11.13.3.6 Diffusion Measurements 425
11.13.3.7 Imaging Experiments 425
11.13.4 Borehole NMR 426
11.13.4.1 The Governing Concept 426
11.13.4.2 Measurement Sequences 427
11.13.5 Surface NMR 427
11.13.5.1 The Governing Concept 427
11.13.5.2 Signal Processing 428
11.13.5.3 1-D Inverse Modeling: Magnetic Resonance Sounding 429
11.13.5.3.1 Calibration 430
11.13.5.3.2 T_1 and T_2 relaxation times 430
11.13.5.3.3 Limitation 431
11.13.5.4 2-D Inverse Modeling: Magnetic Resonance Tomography (MRT) 431
11.13.6 The NMR Link to Rock Properties 432
11.13.6.1 Sensitivity to Porosity 432
11.13.6.2 Sensitivity to Permeability 433
11.13.6.2.1 A single pore 433
11.13.6.2.2 Pore size distribution and multiexponential decay 433
11.13.6.2.3 The link to permeability 433
11.13.6.3 Unsaturated Hydraulic Conductivity: Exploring the Vadose Zone 435
11.13.6.4 Petrophysical Characterization Based on Diffusion Measurements 436
11.13.6.4.1 Surface-to-volume ratio 436
11.13.6.4.2 Pore fluid characterization 436
11.13.7 Case Histories 437
11.13.7.1 Hard Rock Reservoir 437
11.13.7.1.1 Objective 437
11.13.7.1.2 Methodology and results 437
11.13.7.1.3 Discussion 438
11.13.7.2 Sedimentary Aquifer 438
11.13.7.2.1 Objective and methodology 438
11.13.7.2.2 Results and discussion 438
11.13.7.3 Groundwater Storage 439
11.13.7.3.1 Objective 439
11.13.7.3.2 Methodology and results 439
11.13.7.4 Saltwater Intrusion 440
11.13.7.4.1 Objective 440
11.13.7.4.2 Methodology and results 441
11.13.7.4.3 Discussion 441
11.13.7.5 Permafrost 441
11.13.7.5.1 Objective 441

11.13.7.5.2 Methodology and results 442
11.13.7.5.3 Discussion 442
11.13.8 Summary 442
Acknowledgments 443
References 443

Abbreviations

BVI	Bulk volume of irreducible saturation
CPMG	Carr–Purcell–Meiboom–Gill pulse sequence to obtain T_2 relaxation times
FID	Free induction decay
MRS	Magnetic Resonance Sounding
NMR	Nuclear Magnetic Resonance
PFG	Pulsed field gradient
SR	Saturation recovery experiment

Symbols

$\mathrm{B_0}$	Static magnetic field vector
$\boldsymbol{B_0}$	Magnitude of the static magnetic field
$\mathrm{B_1}$	Magnetic field vector rotating at the Larmor frequency
$\mathrm{B_{1\perp}^{+}}$	Circularly polarized with **M** corotating and to $\mathbf{B}_0$ perpendicular component of $\mathbf{B}_1$
$\boldsymbol{D}$	Self-diffusion coefficient
$\boldsymbol{\eta}$	Dynamic viscosity
$\boldsymbol{\gamma}$	Gyromagnetic ratio $= 2.67 \times 10^{-8}$ rad $\mathrm{s^{-1}\,T^{-1}}$ for hydrogen protons
$\boldsymbol{G}$	Magnitude of magnetic field gradients
$\boldsymbol{\hbar}$	Planck's constant divided by $2\pi = 1.055 \times 10^{-34}$ J s
i, j, k	Unit vectors along the *x*-, *y*-, and *z*-axes
$\boldsymbol{k}$	Permeability
$\boldsymbol{K}$	Hydraulic conductivity
$\boldsymbol{k_B}$	Boltzmann's constant $= 1.381 \times 10^{-23}$ J $\mathrm{K^{-1}}$
$\mathrm{M_0}$	Thermal equilibrium magnetization vector
$\boldsymbol{M_0}$	Magnitude of the thermal equilibrium magnetization
M	Magnetization vector
m	Microscopic magnetization vector
$\boldsymbol{M_x}$, $\boldsymbol{M_y}$, $\boldsymbol{M_z}$	Cartesian components of the magnetization vector
$\boldsymbol{N}$	Number of spins per volume $= 6.692 \times 10^{28}$ $\mathrm{m^{-3}}$ for hydrogen protons
$\boldsymbol{\phi}$	Porosity
P90	Pulse that rotates the magnetization by 90°
$\boldsymbol{q}$	Pulse moment (pulse current *I* times pulse duration τ_p)
$\boldsymbol{r_{tube}}$	Radius of a cylindrical tube
r	Spatial location vector
$\boldsymbol{\rho}$	Electrical resistivity
$\boldsymbol{\varrho}$	Surface relaxivity
$\boldsymbol{S/N}$	Signal to noise
$\boldsymbol{S/V}$	Surface area to pore volume
$\boldsymbol{T_1}$	Longitudinal (or spin-lattice) relaxation time
$\boldsymbol{T_{1,B}}$	T_1 bulk fluid relaxation time
$\boldsymbol{T_2}$	Transveral (or spin-spin) relaxation time
$\boldsymbol{T_{2,B}}$	T_2 bulk fluid relaxation time
$\boldsymbol{T_2^*}$	Free induction decay relaxation time
T	Tortuosity
$\boldsymbol{t_E}$	Time between two echoes
$\boldsymbol{\tau_p}$	Pulse duration
$\boldsymbol{\tau}$	Time between pulse and maximum of an echo
$\boldsymbol{\vartheta}$	Temperature
$\boldsymbol{\omega}$	Water content
$\boldsymbol{\omega_L}$	Angular resonant frequency (e.g., the Larmor frequency)

11.13.1 Introduction

The widespread use of nuclear magnetic resonance (NMR) began soon after its discovery in the 1940s (Bloch, 1946; Purcell et al., 1946), with a rapid proliferation resulting in the Nobel Prize in Physics in 1952 for Bloch and Purcell. Some of the applications of NMR have become essential tools of modern science, such as NMR spectroscopy, or are present in daily life, such as magnetic resonance imaging (MRI) used in clinical diagnostics. This widespread use is hardly surprising because NMR captures information from one of the key elements of life, the hydrogen proton.

The potential application of NMR for geosciences was envisaged in the 1960s (Varian, 1962), when early NMR logging tools were developed (Brown and Gamson, 1960) in the framework of oil and gas exploration. Broad acceptance with increasing success was followed by pioneering borehole tool designs in the 1980s (Jackson et al., 1980), accompanied by fundamental research on the petrophysical relations of rock properties and NMR-measured parameters. The community regularly meets at the Magnetic Resonance in Porous Media (MRPM) conferences collecting the latest developments in conference proceedings (e.g., Fantazzini et al., 2010) or special issues (e.g., Song, 2012). Another cornerstone has been the development of a device capable of forcing the excitation of

hydrogen protons in the subsurface and recording their NMR signal at the surface (Semenov, 1987; Semenov et al., 1989). This development introduced NMR to a new society of geophysicists, the community of hydrogeophysicists. This community regularly organizes a workshop on Magnetic Resonance in the Subsurface, and the latest research is summarized in special issues (Valla and Yaramanci, 2002; Yaramanci and Legchenko, 2005; Yaramanci et al., 2008, 2014; Legchenko et al., 2011).

Unlike applying NMR as a spectroscopy tool, the geoscience community uses NMR for imaging and utilizes the processes that determine the relaxation behavior of the measured signal to estimate rock properties. Discovering that rock properties (such as permeability and tortuosity), pore fluid properties (such as diffusion coefficients), and wettability control the relaxation process provided NMR with a high level of importance within the set of geophysical measurement techniques.

In the following sections, we will describe the theoretical framework of NMR (Section 11.13.2) and its realization as laboratory NMR (Section 11.13.3), borehole NMR (Section 11.13.4), and surface NMR (Section 11.13.5). We focus on basic experiments that provide access to the technique and on understanding the effect of the main controlling parameters. Section 11.13.6 links the aforementioned experiments to petrophysical properties. We demonstrate how porosity and permeability are theoretically related to measured quantities and provide brief examples. We hopefully motivate the reader to progress with further reading. Finally, in Section 11.13.7, we discuss field examples covering laboratory NMR, borehole NMR, and surface NMR and ranging from rock permeability to saltwater intrusion and permafrost studies.

11.13.2 Fundamentals of NMR

In this section, we give a brief introduction to the fundamental physics of NMR. We show that a macroscopic magnetization is built up by an ensemble of single nuclear spins, why the orientation of this magnetization can be manipulated by electromagnetic pulses, and how the relaxation process after this manipulation is described.

11.13.2.1 Spin Magnetization

The quantum mechanical property of matter called spin is one of the four basic physical properties of atoms, in addition to mass, electric charge, and magnetism. Because matter is composed of atoms, there are macroscopic consequences, that is, matter has a mass and, depending on its internal structure, may be electrically charged. These two properties are easily observable. The other two properties, nuclear magnetism and nuclear spin, are less evident. Nuclear magnetism is very weak. A high bulk magnetism of some material is due to the electrons, not the nuclei. Nuclear spin, finally, is even less tangible. As an analogy, one may imagine nuclear spin as an atomic nucleus that is rotating in a manner similar to a tiny planet in space but this is a rotational angular momentum. The spin angular momentum, however, is an intrinsic property of the nucleus without any rotation (Levitt, 2002).

There is no simple macroscopic consequence, such as a macroscopic spin; but exposing an ensemble of spins to a static magnetic field $\mathbf{B}_0$ will cause a tiny macroscopic magnetization $\mathbf{M}_0$ based on the microscopic nuclear spins. An exact treatment to calculate $\mathbf{M}_0$ utilizes the principles of quantum mechanics and therefore is beyond the scope of this chapter and can be found elsewhere (e.g., Levitt, 2002 or Callaghan, 2007). Nevertheless, a semiclassical description of the phenomenon is useful.

We associate a single spin with a microscopic magnetization $\mathbf{m}$. Without an external magnetic field and considering a set of spins, no macroscopic magnetization is present because the orientation of all microscopic magnetizations $\mathbf{m}$ is random in space, thus canceling their magnetizations from a macroscopic perspective. Exposing a single spin to a magnetic field $\mathbf{B}$, the rate of change of a magnetization with time is described by

$$\frac{d\mathbf{m}(t)}{\mathrm{d}t} = \gamma[\mathbf{m}(t) \times \mathbf{B}]\cdot \qquad [1]$$

The solution to eqn [1], with the static field $\mathbf{B}_0$

$$\mathbf{B} = \mathbf{B}_0 = 0\mathbf{i} + 0\mathbf{j} + B_0\mathbf{k} \qquad [2]$$

where $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are unit vectors along the x-, y-, and z-axes, is a precession with a frequency, the Larmor frequency,

$$\omega_\mathrm{L} = \gamma \cdot B_0 \qquad [3]$$

at a cone orientated along the direction of the static field. The gyromagnetic ratio γ is specific to the nucleus. Now, again, consider a set of spins. The spatial location of each single $\mathbf{m}$ is not constant due to Brownian motion. Thus, locally, the direction and the intensity of the static field exhibit small fluctuations caused by the nearby microscopic magnetizations. These fluctuations in $\mathbf{B}_0$ cause each single $\mathbf{m}$ to fluctuate. This ongoing change, together with the parallel orientation of $\mathbf{m}$ with the direction of $\mathbf{B}_0$ being more stable than the antiparallel orientation (in terms of quantum mechanics, these are lower and higher energy states), leads to slightly more spins orientated with the static field and a macroscopic magnetization $\mathbf{M}_0$ in thermal equilibrium for a unit volume of water, given as

$$\mathbf{M}_0 = \frac{N\gamma^2\hbar^2}{4k_\mathrm{B}\vartheta}\mathbf{B}_0 \qquad [4]$$

with temperature ϑ, the gyromagnetic ratio γ, Planck's constant $\hbar$, Boltzmann's constant k_B, and the number of spins per unit volume N.

11.13.2.2 Excitation

We found that an ensemble of spins exposed to a static magnetic field $\mathbf{B}_0$ will exhibit a macroscopic magnetic moment $\mathbf{M}_0$ aligned with $\mathbf{B}_0$ after a certain time. While $\mathbf{B}_0$ remains active, now, one can manipulate the orientation of $\mathbf{M}_0$ by additionally applying a second magnetic field $\mathbf{B}_1$ oscillating at ω_L and transverse to $\mathbf{B}_0$. It is common to use $\mathbf{M}_0$ only for the equilibrium state. Whenever manipulating the macroscopic magnetization and calculating a change with time, we refer to $\mathbf{M}$ with $\mathbf{M}$ $(t=0)$ equals $\mathbf{M}_0$. The solution to eqn [1], replacing $\mathbf{m}$ with $\mathbf{M}$ and setting

$$\mathbf{B}(t) = B_1\cos(\omega_\mathrm{L}t)\mathbf{i} - B_1\sin(\omega_\mathrm{L}t)\mathbf{j} + B_0\mathbf{k} \qquad [5]$$

then reads

$$M_x = M_0 \sin(\omega_1 t)\sin(\omega_L t)$$
$$M_y = M_0 \sin(\omega_1 t)\cos(\omega_L t) \quad [6]$$
$$M_z = M_0 \cos(\omega_1 t)$$

where $\omega_1 = \gamma B_1$. Equation [6] implies that on application of an oscillating magnetic field $\mathbf{B}_1$, the magnetization precesses about the direction of $\mathbf{B}_0$ at ω_L and about the 'rotating' direction of $\mathbf{B}_1$ at ω_1. It is useful to introduce a frame rotating with ω_L, that is, with $\mathbf{B}_1$, as shown in Figure 1. Now, it is easy to visualize the impact of $\mathbf{B}_1$ and to understand why it is known as the 'resonant excitation pulse.' Depending on the duration of the pulse $t = \tau_p$, the magnetization will rotate by an angle

$$\beta = \omega_1 \tau_p = \gamma B_1 \tau_p \quad [7]$$

about the direction of $\mathbf{B}_1$.

11.13.2.3 Relaxation

We have forced the spin system away from its thermal equilibrium by applying an excitation pulse at resonant frequency. After this pulse, the system will relax to the equilibrium state as before the pulse. This change of magnetization with time, in the frame rotating with ω_L, is described by the Bloch equation

$$\frac{d\mathbf{M}}{dt} = \gamma[\mathbf{M}(t) \times \mathbf{B}_0] - \frac{M_x \mathbf{i} + M_y \mathbf{j}}{T_2} - \frac{M_z - M_0}{T_1}\mathbf{k} \quad [8]$$

Accordingly, $\mathbf{M}$ still precesses at ω_L about $\mathbf{B}_0$, M_z rebuilds to its equilibrium state with a time constant T_1, and $M_{x,y}$ collapse with a time constant T_2. These time constants T_1 and T_2 are referred to as longitudinal (or spin-lattice) and transveral (or spin-spin) relaxation times, respectively. On the one hand, T_1 relaxation is the process of building up the macroscopic magnetization as described qualitatively in the preceding text and, as the name implies, a process that involves the exchange of energy between the spin system and the surrounding 'lattice' (Callaghan, 2007). Spin–spin relaxation, on the other hand, corresponds to the phase coherence between the nuclear spins.

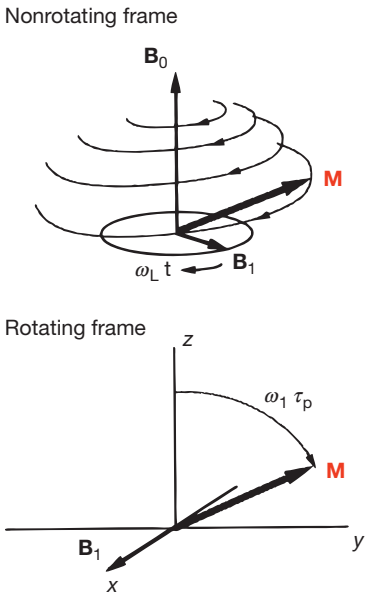


Figure 1 Evolution of the macroscopic magnetization **M** in the presence of a static field $\mathbf{B}_0$ and at the Larmor frequency ω_L oscillating field $\mathbf{B}_1$ in the nonrotating frame and in a frame rotating with ω_L. Redrawn from Callaghan P (2007) Principles of Nuclear Magnetic Resonance Microscopy. Oxford University Press.

11.13.3 Laboratory NMR

This section introduces the basic concepts of measuring NMR at laboratory scale and in a geoscientific context. Because in geoscience, we mainly deal with hydrogen protons that generate the NMR signal, we move from nuclear spins to hydrogen protons in our description. We give the basic design of a laboratory NMR system, present the main sequences, and explain why and for what reason these sequences are used.

11.13.3.1 The Governing Concept

Following the fundamentals of NMR, a device to generate and measure NMR signals needs a static magnetic field to build up a magnetization, a secondary oscillating magnetic field to force excitation, and a detector to record the generated signal.

A very basic design is shown in Figure 2. Here, a permanent magnet is used to provide the static magnetic field $\mathbf{B}_0$. However, iron-core electromagnets or even superconducting magnets are also frequently used (Fukushima and Roeder, 1981). In any case, the key challenge is to ensure a high homogeneity of the field. Typically, the field inhomogeneity is below a few Hz or tens of nT (Callaghan, 2007).

The $\mathbf{B}_1$ field, that is, the excitation pulse, is then generated by a coil. The sample is placed within the coil, and both are placed within the most homogeneous part of the $\mathbf{B}_0$ field. This ensures homogeneity in ω_L and therefore homogeneous excitation throughout the sample. If, however, $\mathbf{B}_0$ is inhomogeneous and

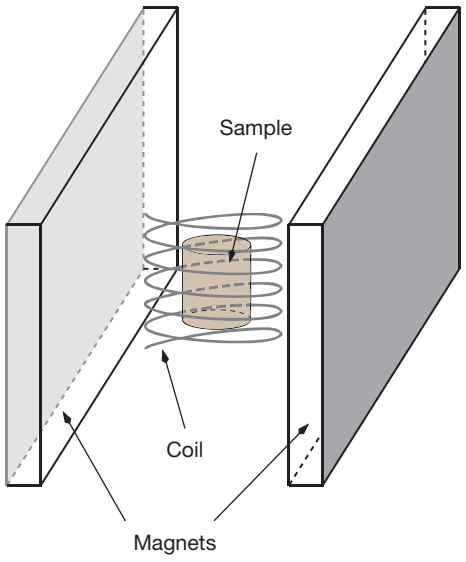


Figure 2 Basic design of a laboratory NMR system. Permanent magnets are used to build the static magnetic field $\mathbf{B}_0$. A multiturn coil is used to generate the $\mathbf{B}_1$ fields for excitation. The same coil is used to detect the NMR signal. The sample to be measured is placed within the coil.

therefore ω_L changes throughout the sample, then the $\mathbf{B}_1$ pulse may become volume-selective since excitation belongs to a match of ω_L and the frequency of $\mathbf{B}_1$. Thus, a well-defined gradient in $\mathbf{B}_0$ opens the possibility of imaging different parts of the sample. Such well-defined gradients may be built by a Maxwell pair or planar coils (Callaghan, 2007). The NMR relaxation signal is finally picked up as a voltage response by the same coil.

11.13.3.2 Free Induction Decay

The free induction decay (FID) experiment is the easiest NMR experiment to be carried out. The experiment consists of emitting $\mathbf{B}_1$ perpendicular to $\mathbf{B}_0$ at ω_L for a duration τ_p, followed by signal detection (see **Figure 3**). Assuming $\mathbf{B}_0$ to be in the z-direction, the magnetization after the pulse is given as (equivalent to eqn [6] but in a frame rotating with ω_L)

$$\begin{aligned} M_z &= M_0 \cdot \cos\beta \\ M_{x,y} &= M_0 \cdot \sin\beta \end{aligned} \qquad [9]$$

In the typical notation, a pulse that rotates the magnetization by 90°, that is, transfers the magnetization from the z-direction to the x, y plane, is referred to as P90. The solution of eqn [8] then yields

$$M_z(t) = M_0 \cdot \left(1 - e^{(-t/T_1)}\right) \qquad [10]$$

$$M_{x,y}(t) = M_0 \cdot e^{(-t/T_2)} \qquad [11]$$

Thus, both relaxation processes are described by an exponential function. But only $M_{xy}(t)$ can be measured as an induced current in a coil, because only $M_{xy}(t)$ is oscillating at ω_L in the nonrotating frame. Consequently, the signal detected after a P90 is then referred to as FID.

According to eqn [4], the initial amplitude of the FID, that is, the signal amplitude at $t=0$, scales linear with the amount of hydrogen protons being excited by the experiment. This relationship builds the base of NMR to directly quantify the amount of water being detected and to calculate the water content of a measured sample.

In contrast to eqn [11], FID signal may be described by

$$\mathrm{FID}(t) = M_0 \cdot e^{(-t/T_2)} \cdot f(t, B_{ih}) \qquad [12]$$

where $f(t, B_{ih})$ is a function that considers any inhomogeneity B_{ih} of the static magnetic field (Grunewald and Knight, 2012). Remember that the macroscopic magnetization is built up from a set of protons. The measured signal then is a synchronous rotation of all of their single spins. If any inhomogeneity in $\mathbf{B}_0$ is present, the Larmor frequency will slightly change within this set of protons. Thus, the synchronous rotation of all spins will become out of phase with time (as depicted as dephasing in **Figure 4**), and therefore, the signal decay is controlled more by a dephasing caused by the magnetic inhomogeneities than by the collapse of $M_{x,y}$ (Grunewald and Knight, 2011). For instance, variations in the magnetic field in the range of tens of nT cause frequency variations of some Hz. These inhomogeneities can be on a macroscale, such as variations in the field strength of $\mathbf{B}_0$, or on a small scale due to particles with a significant magnetic susceptibility in the porous medium. Both variations of the magnetic field intensity lead to slightly different Larmor frequencies and thus to a loss of phase coherence at the microscale. Thus, an experiment that reduces the effect of magnetic gradients is necessary.

11.13.3.3 The Transverse Relaxation Time T_2

It has been reported by Hahn (1950) that a two-pulse sequence consisting of a P90 followed by a pulse that rotates the magnetization by 180° (P180) and separated from the P90 by a time τ generates an echo exactly at time τ after the P180. The key to this experiment is that the P180 inverts the phase relation of the spins accumulated during time τ and due to the rotation with a different Larmor frequency. At a time τ after the P180, the same phase relation has been accumulated again,

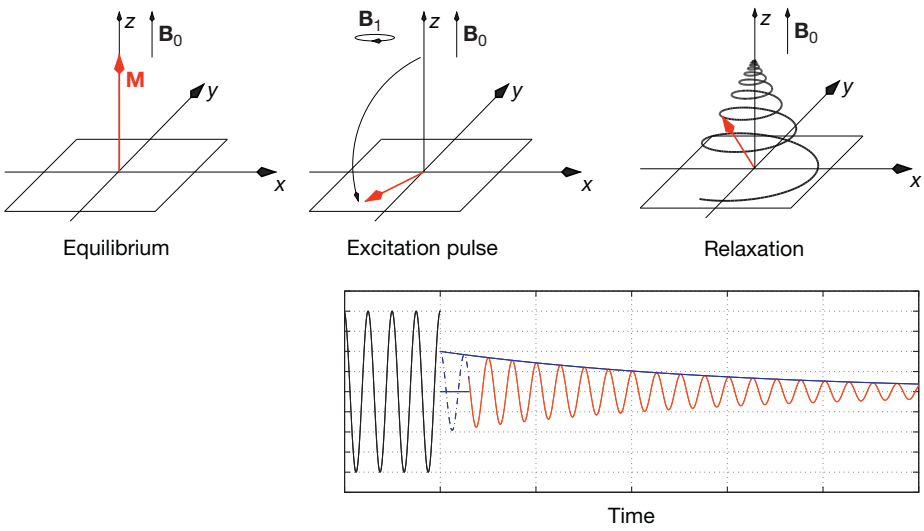


Figure 3 The FID experiment. An excitation pulse $\mathbf{B}_1$ perpendicular to the static magnetic field $\mathbf{B}_0$ rotates the macroscopic magnetization **M** by 90°. The relaxation process of **M** back to equilibrium is observed as FID.

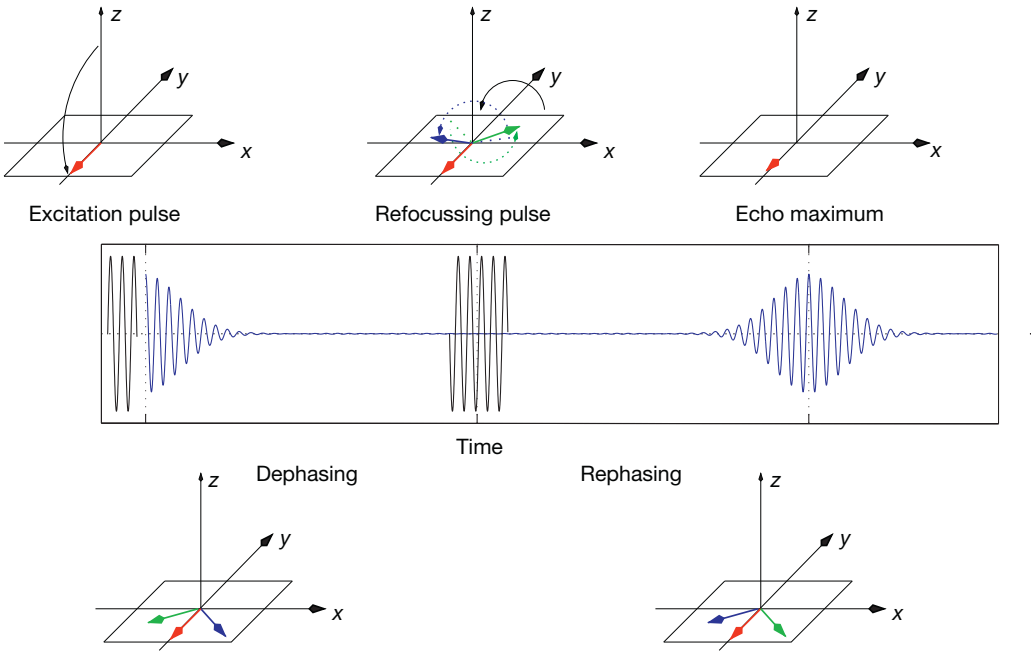


Figure 4 The Hahn echo experiment. The first part of the experiment is the FID, as in **Figure 3**, but is controlled by magnetic field inhomogeneities that cause dephasing. A refocusing pulse rotates the magnetization by 180°, which reverses all the phase relations and leads to a rephasing. Finally, an echo can be observed.

and a refocusing can be observed (due to the previous inversion), referred to as the Hahn echo (**Figure 4**).

The extension of this experiment is the Carr–Purcell–Meiboom–Gill (CPMG) sequence (Carr and Purcell, 1954; Meiboom and Gill, 1958), which consists of one primary P90 and a 'train' of P180 pulses, each separated by a constant time

$$t_E = 2\tau \qquad [13]$$

In between two P180s, an echo will be observed. Together, the maxima of the echoes form the T_2 relaxation curve, as given in eqn [11]. There are two basic pieces of information encoded in this experiment: (i) the initial amplitude M_0 again encoding the amount of detected water and (ii) the transverse relaxation time T_2. This relaxation time T_2 of water in porous media consists of the bulk relaxation $T_{2,B}$, surface relaxation $T_{2,S}$, and relaxation due to molecular diffusion $T_{2,D}$. All of these components are independent processes, and the inverse of T_2 is the sum of the inverse of these components:

$$\frac{1}{T_2} = \frac{1}{T_{2,B}} + \frac{1}{T_{2,S}} + \frac{1}{T_{2,D}} \qquad [14]$$

Relaxation due to diffusion $T_{2,D}$ occurs if a proton moves through either a macroscopic or a microscopic magnetic gradient. While moving, the proton passes through regions with varying magnetic field strength, that is, it not only experiences a single Larmor frequency in between two P180s but also accumulates changes due to the varying Larmor frequency. Because of the different pathways of single protons within a set of protons, a perfect phase inversion does not exist, and a certain dephasing remains. This diffusion effect is non-removable. The diffusion term can be expressed via eqn [15] (e.g., Kleinberg and Horsfield, 1990)

$$\frac{1}{T_{2,D}} = D\frac{(\gamma G t_E/2)^2}{3} \qquad [15]$$

where G is the magnetic field gradient and D is the self-diffusion coefficient of the fluid, that is, the mobility of the protons in the fluid. According to eqn [15], diffusion effects can be minimized by selecting a small t_E while conducting a CPMG sequence. If the diffusion term can be assumed to be negligible, only $T_{2,B}$ and $T_{2,S}$ affect the relaxation time in eqn [14].

The bulk relaxation term $T_{2,B}$ denotes the relaxation time of a free fluid in a perfectly homogeneous magnetic field B_0. Among other things, T_B is controlled by the self-diffusion of the fluid (Bloembergen et al., 1948). For a particle with a hydrodynamic radius a approximately equal to that of the medium, the Stokes–Einstein equation is used as an approximation:

$$D = \frac{k_B \vartheta}{6\pi a \eta} \qquad [16]$$

where k_B is the Boltzmann's constant, ϑ is the absolute temperature, and η is the dynamic viscosity (Krynicki et al., 1978). Because the concentration of dissolved oxygen and ferromagnetic and paramagnetic ions (such as iron or manganese) can further reduce the measured relaxation time compared to uncontaminated water (Bloembergen et al., 1948), exact values of $T_{2,B}$ should be estimated using separate measurements on the bulk fluid. However, the temperature dependency of $T_{2,B}$ is still mainly controlled by D and can be described as (Bloembergen et al., 1948; Simson and Carr, 1958)

$$T_B \propto \frac{\vartheta}{\eta} \qquad [17]$$

11.13.3.4 The Longitudinal Relaxation Time T_1

The application of CPMGs to measure T_2 relaxation times has been of great importance to shortening the duration of NMR experiments. A more time-consuming approach is to perform T_1 experiments to determine the longitudinal relaxation. We have expressed in the preceding text that a change in M_z, and therefore T_1 (eqn [10]), cannot be measured directly because a precise measurement of static magnetic fields would be necessary. However, an indirect measurement is possible through FID experiments. These types of experiments are referred to as inversion recovery (IR) or saturation recovery (SR) experiments and consist of a set of double-pulse experiments.

SR utilizes a P90 to rotate M_z, similar to the FID experiment depicted in Figure 3. The initial signal amplitude is M_0. According to eqn [10], M_z is rebuilt from zero. After a time τ, a second P90 is applied that rerotates M_z by 90°; the signal amplitude depends on the time τ between the two pulses, that is, the amplitude that has been rebuilt during the time τ is observed. By repeating this double-pulse experiment with different τ, the reconstruction of M_z is observed.

The IR uses P180 as the first pulse instead of P90, that is, the reconstruction of M_0 from $-M_0$ can be observed. While T_2 is obtained from a single CPMG that contains thousands of data points, each data point of the T_1 relaxation curve is measured separately.

11.13.3.5 Inverse Modeling: From Measured Data to Relaxation Times

The aforementioned paragraphs have described the sequences to measure the data containing T_2 and T_1 relaxation times. It has also been mentioned which processes impact on those relaxation times. The missing step in between, the estimation of T_2 or T_1 from the measured data is done by inverse modeling. This step is often referred to as the inverse Laplace transformation. The following explanation refers to T_2 but, with slight changes, is valid for T_1 as well.

In a simple case, the measured signal may contain one relaxation time, that is, the signal decays monoexponentially. Using the logarithm of the data, a relaxation time T_2 and magnetization M_0 (eqn [11]) can be easily obtained by applying a linear regression.

The more general case treats a multiexponential signal. The measured data d (being voltages in the detection coil) are described by a sum of N relaxation times T_2

$$d(t) = M_{x,y}(t) = \sum_{i=1}^{N} A_i \cdot e^{-1/T_{2,i}} \qquad (18)$$

with A_i an amplitude (in terms of voltages) associated with the 'ith' relaxation time. Commonly, a minimum structure or smooth distribution of relaxation times (Whittall et al., 1991) is used (Figure 5). Thus, a number of relaxation times are fixed, mostly in a logarithmic manner, and the corresponding magnetization is estimated using algorithms based on least-squares optimization. A regularization parameter is used to weight between the minimum structure in the T_2 distribution (i.e., the distribution may not rapidly change) and the minimum residual between the measured data and model response.

11.13.3.6 Diffusion Measurements

Although they are extensively used, the aforementioned experiments are only a small portion of experiments used in the geologic framework, commonly referred to as relaxometry. Another field includes experiments that explore the effect of diffusion on NMR experiments. We have already mentioned the fundamental effect of diffusion in eqn [15]. A set of strategies that are used to measure the self-diffusion coefficients of the pore fluid are also available. The first experiments were published by Stejskal and Tanner (1965). The authors used a pulsed field gradient (PFG) to enhance the effect of diffusion on the amplitude of a spin echo. Hürlimann et al. (2002) presented an experiment to determine the diffusion coefficient in a constant gradient field. In contrast to the pulsed gradients that are applied within a certain time interval in between two excitation pulses, the constant field is permanent during the complete sequence of pulses. The third approach was presented by Song (2003) in which the natural internal gradients inside a pore are used. Due to the success of these experiments and the complementary information that can be obtained from relaxation times and diffusion coefficients, both types of experiments can be combined to create multidimensional experiments, as described by Hürlimann and Venkataramanan (2002).

11.13.3.7 Imaging Experiments

The ability of laboratory NMR devices to create gradients in the $\mathbf{B}_0$ field does not only allow to explore diffusion effects on relaxation times but also, probably much more well known in medical application, allow for a spatial localization of the NMR signal. In Section 11.13.3.1, we have mentioned that

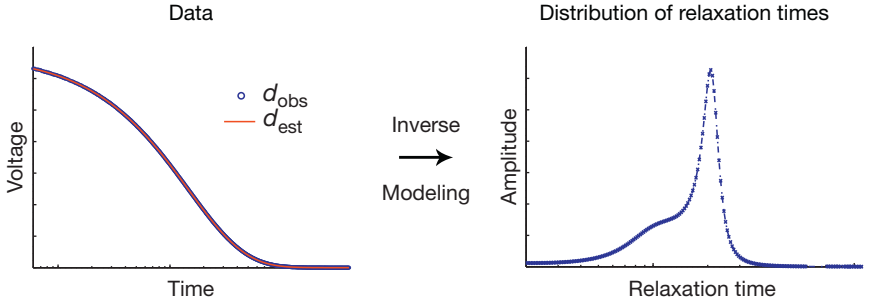


Figure 5 Inverse modeling of CPMG measured data to derive a smooth distribution of amplitudes corresponding to a number of T_2 relaxation times by minimizing the residual between measured d_{obs} and model response d_{est}.

gradients in $\mathbf{B}_0$ lead to gradients in ω_L that are used for imaging. Two strategies to derive images may be distinguished.

Firstly, we are targeting a broad frequency spectrum of the $\mathbf{B}_1$ excitation pulse. Since excitation belongs to a match of ω_L and the frequency of the $\mathbf{B}_1$ pulse, excitation throughout a complete sample can be derived by controlling the shape and duration of a pulse according the gradient of $\mathbf{B}_0$. The spatial location of a specific signal within the sample is then encoded by its specific ω_L and can be decoded simply by the Fourier transformation (see Figure 6).

Secondly, one can use selective excitation. In this case, the $\mathbf{B}_0$ gradient and the $\mathbf{B}_1$ pulse parameter are chosen in a way that only protons in a certain area are stimulated. For instance, a pulse with a duration in the order of 1 ms will stimulate protons within approximately 1 kHz bandwidth. Changing the center frequency of the pulse then leads to spatial resolution. Obviously, both strategies may also be combined, and further imaging methods exist that are discussed in a number of reviews (e.g., Mansfield and Morris, 1982 or Callaghan, 2007).

11.13.4 Borehole NMR

This section introduces the basic concepts in taking advantage of NMR in borehole application. Due to the diversity in design of NMR tools developed over the last decades, depending on the envisaged target, this review is not complete. Therefore, we present the basic idea and point out the differences to laboratory NMR.

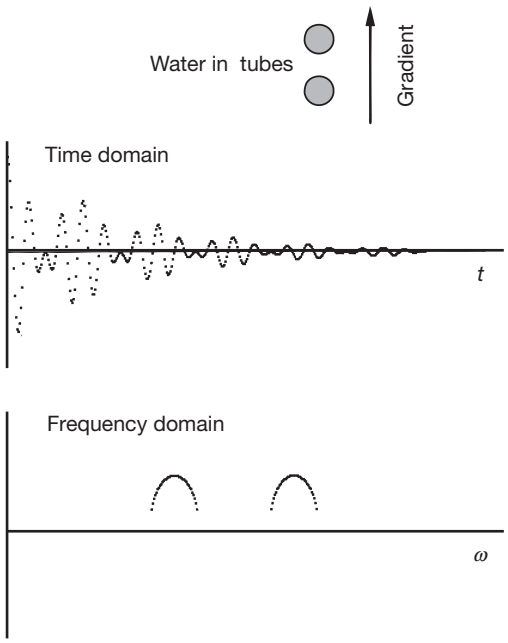


Figure 6 Sketch of 1-D imaging by means of Fourier transformation. Under the presence of gradient in $\mathbf{B}_0$, the spatial location of a signal is encoded in ω_L and can be decoded using Fourier transformation. Redrawn from Callaghan P (2007) Principles of Nuclear Magnetic Resonance Microscopy. Oxford University Press.

11.13.4.1 The Governing Concept

The first NMR logging tools (Brown and Gamson, 1960) were designed to operate with the Earth's magnetic field as $\mathbf{B}_0$ field. A coil was used to emit the $\mathbf{B}_1$ excitation pulse. This setup leads to an excitation throughout the complete surrounding volume with a distribution of excitation angles being inhomogeneous. Thus, not only the hydrogen protons within the surrounding formation but also the ones within the borehole fluid and a potentially present mud cake will contribute to a measured signal.

To overcome this disadvantage, new designs (Cooper and Jackson, 1980; Jackson et al., 1980; Kleinberg et al., 1992) were introduced that utilize permanent magnets to build up an artificial $\mathbf{B}_0$. Obviously, the magnetic field strength decreases with distance to the tool, and therefore, ω_L changes throughout the formation to be investigated. This, however, enables a selective excitation depending on the frequency of $\mathbf{B}_1$, similar to the imaging concept in laboratory NMR. Consequently, modern borehole NMR tools are insensitive in the area close to the tool but capture signals arising from a volume of a certain thickness and distance to the borehole tool. A schematic drawing close to the design of Jackson et al. (1980) is depicted in Figure 7. The basic trade-off to deal with is the distance of the sensitive volume to the tool versus the signal-to-noise ratio. While it is preferable to capture signals from far inside the formation, the signal amplitude is proportional to the amplitude of $\mathbf{B}_0$ (eqn [4]) and thus decreases with increasing distance. Typical frequencies are in the range of hundreds of kHz, while penetration into the formation is about few inches.

The aforementioned tools were wire-line tools, logging the formation after drilling. In the meantime, logging-while-drilling tools have been developed (Menger and Sands, 2003) providing NMR logs during the drilling progress. Typically, tools as used in oil and gas exploration have large diameter (>5 in.) and are several meters long. In the context of groundwater investigations, slim tools that fit into groundwater monitoring wells were recently presented (Walsh et al., 2013).

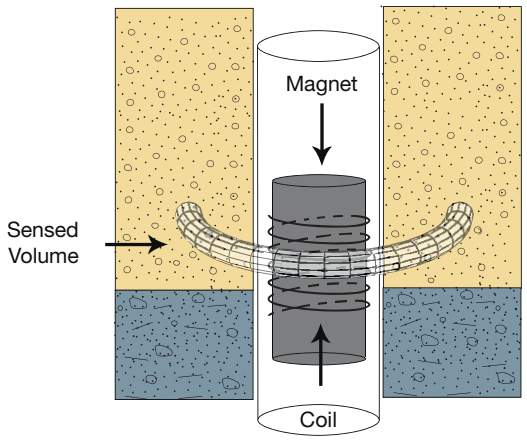


Figure 7 Basic design of a borehole NMR tool. There are various other designs for magnets and coil, but all share the principle of a permanent magnet building the static magnetic field $\mathbf{B}_0$ and a coil used for emitting the $\mathbf{B}_1$ excitation field and for detection of NMR signals. The location of the sensed volume is then defined by the match of excitation and Larmor frequency.

11.13.4.2 Measurement Sequences

A key challenge for the application of NMR in boreholes is to understand the spin dynamics and therefore the applicability of pulse sequences such as a CPMG (to measure T_2 relaxation) in inhomogeneous magnetic fields (magnetic gradients of several Gauss per cm are common) as they come along with those tools. On the one hand, these gradients impact on T_2 by enhanced diffusion (eqn [15]). On the other hand, off-resonant excitation has to be considered. It has been reported (Dunn et al., 2002) that in some cases, the first echoes of a CPMG sequences are seriously influenced, while the influence is diminishing on subsequent echoes. But, for the case of grossly inhomogeneous fields, the initial relaxation rate may also become a weighted rate of T_1 and T_2 (Hürlimann and Griffin, 2000).

In contrast to laboratory NMR, PFG sequences used to obtain diffusion coefficients are not available, but still, diffusion can be addressed by changing the interecho spacing t_E in a CPMG sequence or using combined T_1/T_2 sequences as proposed by Hürlimann and Venkataramanan (2002).

While wire-line tools usually do not measure T_1 relaxation time since the measurement scheme is time-consuming, logging-while-drilling tools measure T_1 based on the SR sequence. Prammer et al. (2000) reported that the SR sequence is even preferable compared to CPMG since it can be applied in a less motion-sensitive manner. This is of great importance taking into account the vibration of the drilling process.

11.13.5 Surface NMR

This section gives the adaption of NMR to nondestructively explore the subsurface using a technique placed at the Earth's surface. We explain the governing concept that is different in many details from laboratory and borehole NMR since the propagation of electromagnetic fields needs to be considered in detail. Furthermore, the state-of-art signal processing, inversion techniques, and limitations of surface NMR are discussed.

11.13.5.1 The Governing Concept

Surface NMR fuses NMR with field-scale electromagnetics to explore the subsurface. The overall concept is depicted in Figure 8. Surface NMR operates in the Earth's magnetic field instead of using artificial magnetic fields, resulting in a Larmor frequency ω_L of 0.9–3 kHz and the macroscopic magnetization $\mathbf{M}$ to line up with the direction of the Earth's field $\mathbf{B}_0$. A large surface loop, commonly circular or rectangular in shape and with a side length ranging from a few meters to 150 m, is used to emit the excitation field $\mathbf{B}_1$ and to receive the NMR signal.

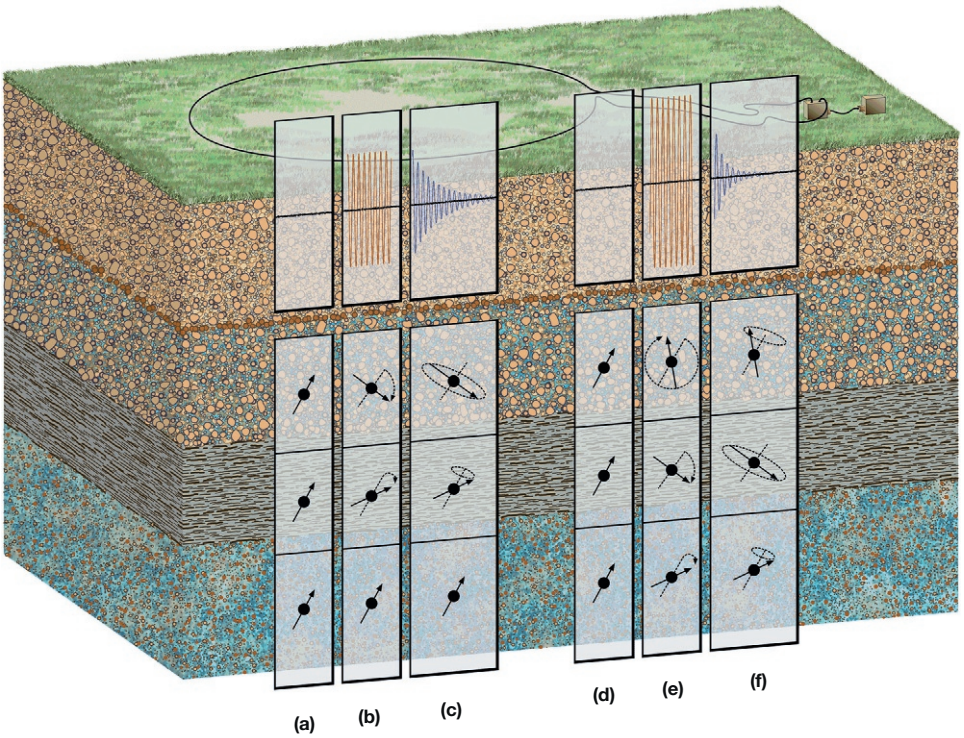


Figure 8 Schematic representation of the magnetic resonance sounding (MRS) experiment. An MRS experiment consists of several FID experiments under the Earth's magnetic field condition with excitation pulses emitted by a large surface loop. In equilibrium (a), all spins are orientated along the direction of the Earth's magnetic field. An excitation pulse (b) causes a distribution of flip angles in the subsurface depending on the amplitude and duration of the pulse. The relaxation recorded (c) is then a superposition of all single relaxations in the subsurface weighted by their respective flip angles and a depth-dependent receiver sensitivity. The next pulse is separated from its predecessor by a certain time to allow for rebuilding the equilibrium magnetization (d). By changing the pulse amplitude (e), the distribution of flip angles and, thus, the composition of the recorded signal are changed (f). Using several different excitation pulses then allows for estimating subsurface properties from the set of measured signals.

This configuration is referred to as coincident loop configuration. Based on inverse modeling, the spatial distribution of water content and relaxation times can be estimated from the measured surface NMR signals.

The surface NMR conditions of a homogeneous $\mathbf{B}_0$ (to the extent of a homogeneous Earth's field), but the inhomogeneous $\mathbf{B}_1$ distinguishes surface NMR not only from laboratory NMR but also from borehole NMR. Laboratory NMR utilizes homogeneous $\mathbf{B}_0$ and $\mathbf{B}_1$ fields and therefore provides homogeneous excitation angles throughout a sample. In the case of borehole NMR, both $\mathbf{B}_0$ and $\mathbf{B}_1$ are inhomogeneous. Consequently, ω_L is inhomogeneous and matches the frequency of $\mathbf{B}_1$ only in a thin area.

Concerning surface NMR ω_L is homogeneous and therefore the frequency of $\mathbf{B}_1$ can match ω_L throughout the complete subsurface. But due to the inhomogeneous $\mathbf{B}_1$, the excitation angles are in general inhomogeneous. A distribution of excitation angles is present and reads

$$\beta(q, \mathbf{r}) = \left|\gamma q \mathbf{B}_{1\perp}^{+}(\mathbf{r}, \rho)\right| \quad [19]$$

where $\mathbf{B}_{1\perp}^{-}(\mathbf{r})$ is the amplitude of the circular with **M** corotating component of $\mathbf{B}_{1\perp}$ $(\mathbf{r})$ (the component of $\mathbf{B}_1$ perpendicular to $\mathbf{B}_0$) normalized to a unit current, **r** the spatial location, γ the gyromagnetic ratio, ρ the resistivity in the subsurface, and q the pulse moment as the product of the current I and the pulse duration τ_{p1}. For a detailed derivation and more detailed explanation on the decomposition of $\mathbf{B}_1$ in co- and counterrotating parts, we refer to Weichman et al. (2000) or Hertrich (2008). The decomposition of $\mathbf{B}_1$ becomes necessary because $\mathbf{B}_1$ is elliptically polarized in the presence of a conductive media and due to the Earth's field inclination not being perpendicular to $\mathbf{B}_1$. In many cases, surface NMR utilizes a single excitation pulse and records the FID. The induced FID voltage signal for the coincident loop configuration then reads

$$V(q, t) = -2\omega_L M_0 \int \sin\beta(q, \mathbf{r}) \cdot \left|\mathbf{B}_{1\perp}^{-}(\mathbf{r}, \rho)\right| e^{i2\zeta(r, \omega_L)} \cdot w\left(\mathbf{r}, T_2^*\right) \cdot e^{-t/T_2^*} \cdot d^3r \cdot dT_2^* \quad [20]$$

where $\mathbf{B}_{1\perp}^{-}(\mathbf{r})$ is the amplitude of the circular with **M** counterrotating component of $\mathbf{B}_{1\perp}$ $(\mathbf{r})$ normalized to a unit current, ζ is a phase factor important for the decomposition of the elliptically polarized field into co- and counterrotating components, and each subsurface volume element d^3r is described by its water content $w(\mathbf{r}, T_2^*)$; that is, the total water content of a volume element is split into partial water contents each associated with a relaxation time T_2^*. Water content is derived from magnetization via $|\mathbf{M}(\mathbf{r})| = M_0 \cdot w(\mathbf{r}, T_2^*)$. We like to emphasize that in our notation, M_0 needs to be calculated accounting for the two hydrogen protons; otherwise, eqn [20] needs to be multiplied by 2.

There are two extensions to eqn [20]. One is the use of a separated transmitter and receiver loops (Hertrich et al., 2005) that result in the imaging of two-dimensional distributions in high resolution (Hertrich et al., 2009). In this case, the geometric relation between the transmitter and receiver needs to be considered. The other extension is to include double-pulse experiments to measure T_1 relaxation. For the case of on-resonance, that is, exact matching of the transmitter frequency and Larmor frequency, Walbrecker et al. (2011b) provided the magnetization after the second pulse and for the phase-cycled pseudosaturation recovery sequence (pcPSR). Note that off-resonance conditions (i.e., transmitter frequency and Larmor frequency do not match) have several impacts on the application of surface NMR as shown by Walbrecker et al. (2011a).

As shown in **Figure 8** and eqn [19], a surface loop induces a distribution of excitation angles depending on q. Consequently, a change in q, either in I or in τ_{p1}, leads to another distribution and thus another detected signal. This dependency allows for estimation of the spatial distribution of $w(\mathbf{r}, T_2^*)$ in the subsurface based on inverse modeling of the detected signals as a function of q and t. However, before estimating $w(\mathbf{r}, T_2^*)$ from surface NMR measurements, a certain amount of signal processing is necessary to separate the NMR signal from the EM noise that contaminates the signal.

11.13.5.2 Signal Processing

The signal processing of surface NMR records has become an important field of research in the last few years. Due to the latest generation of instruments (Dlugosch et al., 2011) allowing for the simultaneous acquisition of densely sampled time series by multichannel receivers, the signal-to-noise ratio (S/N) has been significantly improved (Walsh, 2008). However, both the time series and the multichannel receiver must be treated carefully to extract the NMR signal. An appropriate scheme includes the following steps:

1. Despiking → delete short pulse-like events
2. Harmonic noise cancelation (HNC) → multichannel-based subtraction of harmonic signals
3. Demodulation → detection of the complex signal envelope
4. Filter → low-pass filter to improve the S/N
5. Stacking → average of repeated measurements to improve the S/N

The first two steps belong to the treatment of EM noise. In a first-order approximation, we distinguish between two main types of EM noise: (i) short pulse-like events with a duration of a few milliseconds (i.e., spikes) and (ii) harmonic noise (Dalgaard et al., 2012). The former may be due to thunderstorms or electrical fences, while the latter are often caused by power lines. However, a detailed understanding of the sources contaminating the NMR signal, their frequency content, and the time dependency is needed in the future.

The common procedure for despiking is to delete the part of the record that is affected (Dalgaard et al., 2012; Jiang et al., 2011) based on statistical properties. Despiking is essential to apply HNC. Some HNC strategies have been published, and all of these strategies share the use of a detection loop (Tx) to record the EM noise-contaminated NMR response and one or more reference loops (Nx) that simultaneously record at a remote position only EM noise that does not contain an NMR signal. Dlugosch et al. (2011) commented on the need to place Nx at least three times the loop diameter (center to center) apart from Tx to avoid canceling the NMR signal. The essential step in HNC is to obtain a transfer function (TF) that correlates Nx with Tx. Walsh (2008) and Dalgaard et al. (2012) used a time-domain approach based on adaptive or Wiener filtering, while Radic

(2006) and Müller-Petke and Costabel (2013) described a filter in the frequency domain. Müller-Petke and Costabel (2013), comparing time-domain and frequency-domain approaches, summarized that frequency-domain approaches are superior to the time-domain approaches.

The HNC is followed by the demodulation (or quadrature detection as in Fukushima and Roeder, 1981) to detect the complex signal envelope of the NMR signal instead of handling the harmonic signal oscillating with ω_L. This demodulation allows for the application of a low-pass filter to improve the S/N and is needed for downsampling. Special care must be taken not to disturb the NMR signal using too low corner frequency and to maintain the error statistics (Müller-Petke et al., 2011a). Finally, repeated measurements are averaged to further improve S/N. **Figure 9** shows an example of processing raw data.

11.13.5.3 1-D Inverse Modeling: Magnetic Resonance Sounding

According to eqn [20], the surface NMR forward problem is linear in the water content. The measured voltages $V(q, t)$ are obtained by a multiplication of the water content $w(\mathbf{r}, T_2^*)$ as a function of relaxation time and spatial location with a sensitivity kernel

$$\mathbf{G} = -2\omega_L M_0 \cdot \sin\beta(q, \mathbf{r}) \cdot \left|\mathbf{B}_{1\perp}^{-}(\mathbf{r}, \rho)\right| e^{i2\zeta(\mathbf{r}, \omega_L)} e^{-t/T_2^*} \quad [21]$$

Accordingly, $\mathbf{G}$ is a function of the Earth's magnetic field (intensity, inclination, and declination). It is controlled by the loop size d and subsurface resistivity ρ because both parameters affect the B-field propagation. A valid assumption for many applications concerning hydrogeologic problems is a 1-D layered Earth. Thus, we simplify the spatial location $\mathbf{r}$ to a depth vector z, reducing the volume integral in eqn [20] to an integral over z instead of d^3r. This simplification reduces the sensitivity kernel $\mathbf{G}$ to a function of q and z. Next, we discretize eqn [20] to express it as a matrix multiplication:

$$\mathbf{V} = \mathbf{G} \cdot \mathbf{w} \quad [22]$$

Equation [22] is then the basis used for forward and inverse modeling, that is, to transfer the subsurface water content distribution into synthetic data and to obtain the water content from the measured data. **Figure 10** illustrates the concept of this 1-D depth sounding, while **Figure 11** shows $\mathbf{G}$ for a few variations of its controlling parameters.

The first inverse modeling (or imaging) approach (Legchenko and Shushakov, 1998) reduced $V(q, t)$ to its initial values at $t = 0$ using a monoexponential fit and only estimated the total water content distribution over depth. This type of inversion is referred to as initial value inversion. Mueller-Petke and Yaramanci (2010) published the current state-of-the-art inverse modeling formulation accounting for a complete dataset, that is, voltages as a function of q and t, and thus, refer to this inversion as QT inversion. Although the inverse problem is linear, there are some useful transformations that transform the problem into a nonlinear problem. One may assume monoexponential (Günther and Müller-Petke, 2012) or stretched exponential (Behroozmand et al., 2012) T_2^* relaxation time behavior in a single layer, which increases the stability of QT inversion due to reduced degrees of freedom in the model space. The water content solution should be positive and must not extend a certain upper boundary. These limits can be derived by application of a logarithm or scaled tangent. Although $V(q, t)$ is a complex value, amplitudes are typically used. Thus, a transformation of eqn [22] to amplitudes is necessary. Both transformations are given in Mueller-Petke and Yaramanci (2010). Another useful transformation is rotating the data, that is, adding an additional phase until the imaginary part is minimal. Rotated data are used instead of absolute values because absolute values cannot become negative; therefore, the error statistics are changed, and artifacts can be introduced. This procedure is common in laboratory NMR. Nevertheless, we would like to emphasize that for exact data rotation, the phase must be constant over time (or at least linear in the case of off-resonance excitation). Otherwise, some data remain in the imaginary part. Unfortunately, surface NMR data may violate this assumption if a record consists of signals originating from different depths with different relaxation times in an electrically conductive environment.

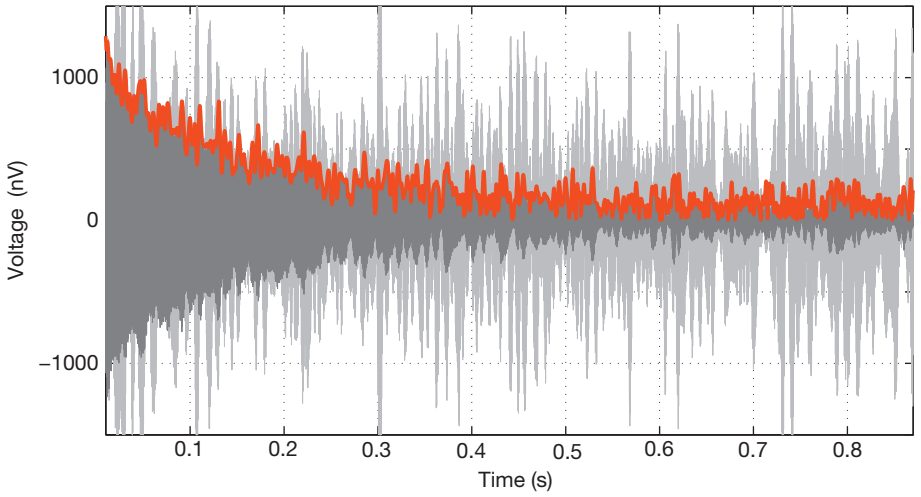


Figure 9 Surface NMR detected signal: raw (light gray) data, after harmonic noise cancelation (dark gray), and after demodulation (red).

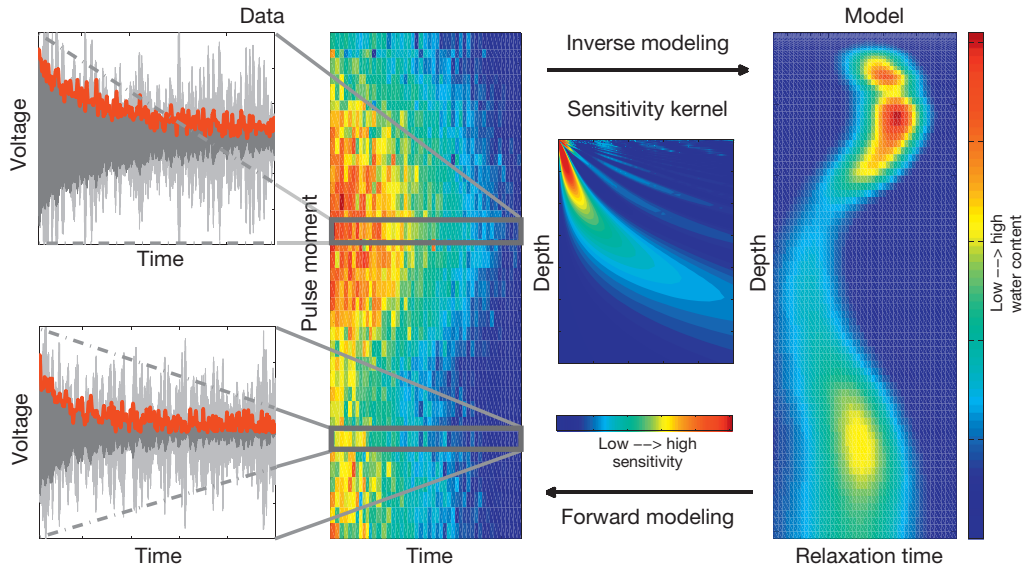


Figure 10 From field data to subsurface models. The MRS dataset consists of all recorded FID measurements for each pulse moment. Inverse modeling transfers the dataset into a subsurface model based on the sensitivity kernel. The most general model is a T_2^* distribution for each depth.

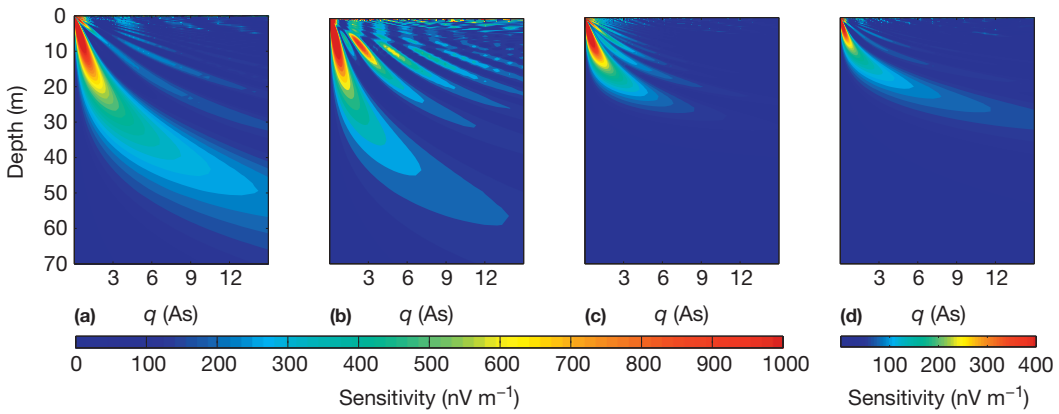


Figure 11 Effect of changing key parameters on the MRS sensitivity kernel. (a) 50 m diameter circular loop, neglecting resistivity, the Earth's field at 48,000 nT and 60° inclination; (b) same as (a) but 10° inclination; (c) same as (a) but conductive half-space of 1 Ω m; (d) same as (a) but 20 m diameter circular loop.

Therefore, using complex data would be preferable, but it is rarely used.

11.13.5.3.1 Calibration

While laboratory and borehole NMR use calibration to calculate water content from measured voltages (see Section 11.13.6.1), a simple calibration is not possible for surface NMR because many factors, such as loop size, resistivity structure, and the Earth's magnetic field properties, affect the measured voltages. In eqn [20], a forward model is provided that allows the water content to be estimated from the measured data using inverse modeling. However, one might ask to which extent the forward model is comprehensive and valid to obtain the correct subsurface water content without calibration. The first field experiment to prove the validity of this forward model was published by Schirov et al. (1991), providing measurements on top of a frozen river. Figure 12 presents the results of a second experiment using the current type of instrumentation that was also conducted on top of a frozen lake and published by Müller-Petke et al. (2011a). This experiment not only demonstrated that the water content of the lake is estimated correctly within the uncertainty of the method but also highlighted the capabilities of resolving the subsurface structure. The boundaries for the frozen ice on top and the bedrock below are fairly well imaged. Note that the water content of the frozen ice is undetectable for surface NMR because the relaxation process is too fast to be measured. The information on ice thickness and depth to bedrock is measured in a small ice drilling and using a hand lead.

11.13.5.3.2 T_1 and T_2 relaxation times

Relaxation times other than T_2^* are of high interest for predicting hydraulic conductivity. As for laboratory NMR, at least double-pulse experiments are required. The first approach to

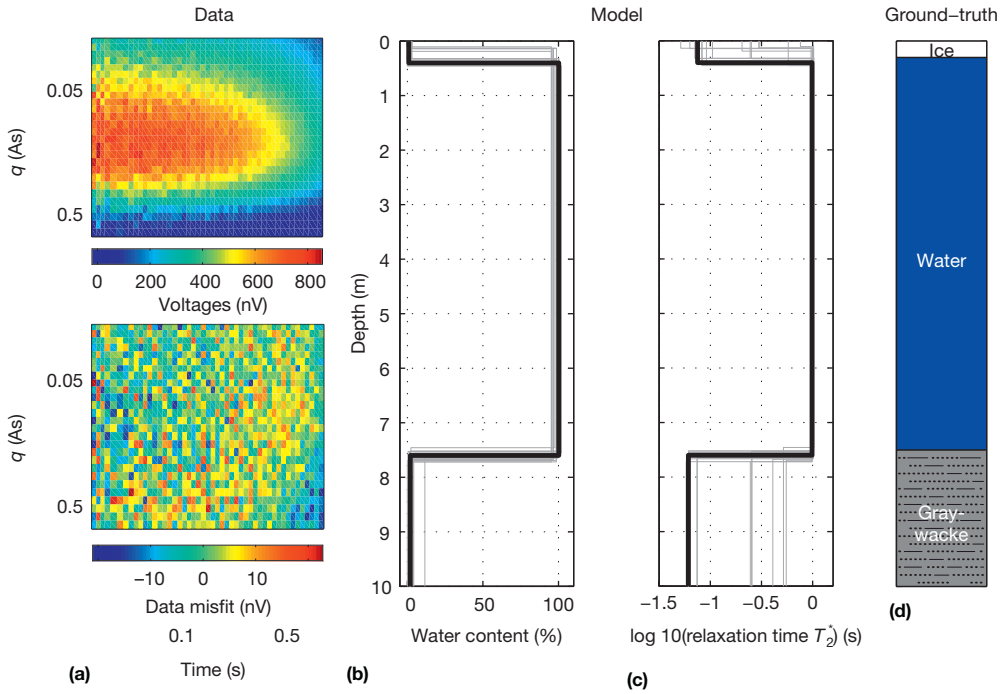


Figure 12 Surface NMR field measurements on top of a frozen lake. The subsurface model in (b) and (c) estimated from the measured data in (a) not only matches the expected water content for free water in a lake but also agrees well with the overall subsurface structure determined from additional measurements. Redrawn from Müller-Petke M, Dlugosch R, and Yaramanci U (2011) Evaluation of surface nuclear magnetic resonance-estimated subsurface water content. New Journal of Physics 13, 095002.

obtain T_1 using a SR-like experiment was published by Legchenko et al. (2004) but was demonstrated to be incapable to suppress unwanted T_2^* components and affected by off-resonance effects (Walbrecker et al., 2011a). Müller-Petke et al. (2013) presented the first inversion scheme to estimate T_1 relaxation time using a pcPSR (Walbrecker et al., 2011b) measurement scheme that minimized off-resonance effects and suppressed unwanted T_2^* components. Note that the inversion of T_1 is a nonlinear problem because T_1 is part of the kernel function. The latest developments also focus on implementing T_2 estimates using spin echo (Legchenko et al., 2010; Grunewald and Walsh, 2013; Grunewald et al., 2014).

11.13.5.3.3 Limitation

Only a few studies have addressed the use of complex data. Braun et al. (2005) demonstrated that the resolution and penetration depth can be improved but mainly used synthetic data and reported on additional phase behavior affecting field data. Nevertheless, Braun and Yaramanci (2008) were able to estimate resistivity from complex surface NMR data under favorable conditions; in this case, favorable conditions include not only an excellent S/N but also reliable phases. These additional phase effects limit the success of complex inversion for field data. Legchenko (2004) and later Walbrecker et al. (2011a) described additional phases due to off-resonance excitation. While the field of complex inversion is still under discussion, there is an ongoing effort to allow for complex inversion in the future.

Unlike laboratory NMR transmitting pulse of few μs, surface NMR emits 5–60 ms excitation pulses. Walbrecker et al. (2009) demonstrated that these long pulses have a significant effect on the recorded FID amplitudes due to relaxation during pulse (RDP). If RDP is neglected, the water content estimated from the FID amplitudes is underestimated. To correct for RDP, an additional extrapolation of half of the pulse duration is proposed. Furthermore, RDP has a strong effect on one of the most significant limitations of surface NMR: the detection of a fast-decaying signal. Although significant reductions in the instrument dead time down to a few milliseconds (Walsh et al., 2011) have been achieved, bound water (as present in clay) cannot be detected because pulses with durations of tens of milliseconds are common.

11.13.5.4 2-D Inverse Modeling: Magnetic Resonance Tomography (MRT)

Many hydrogeophysical problems are sufficiently described by a 1-D layered Earth, at least at the lateral extent of the sensed volume of a magnetic resonance sounding (MRS) measurement. However, variations within the sensed volume still occur, for instance, in the exploration of paleochannel structures that are of great importance in areas affected by glacification. In such cases, imaging the lateral variation of subsurface water content is necessary. A straightforward approach is to conduct several coincident loop soundings at different locations along a profile and to adapt eqn [20] to consider lateral variations (Hertrich et al., 2007). High-resolution images can be obtained, including measurements with separated transmitter and receiver loops as presented by Hertrich et al. (2009). A coincident loop allows changes to be

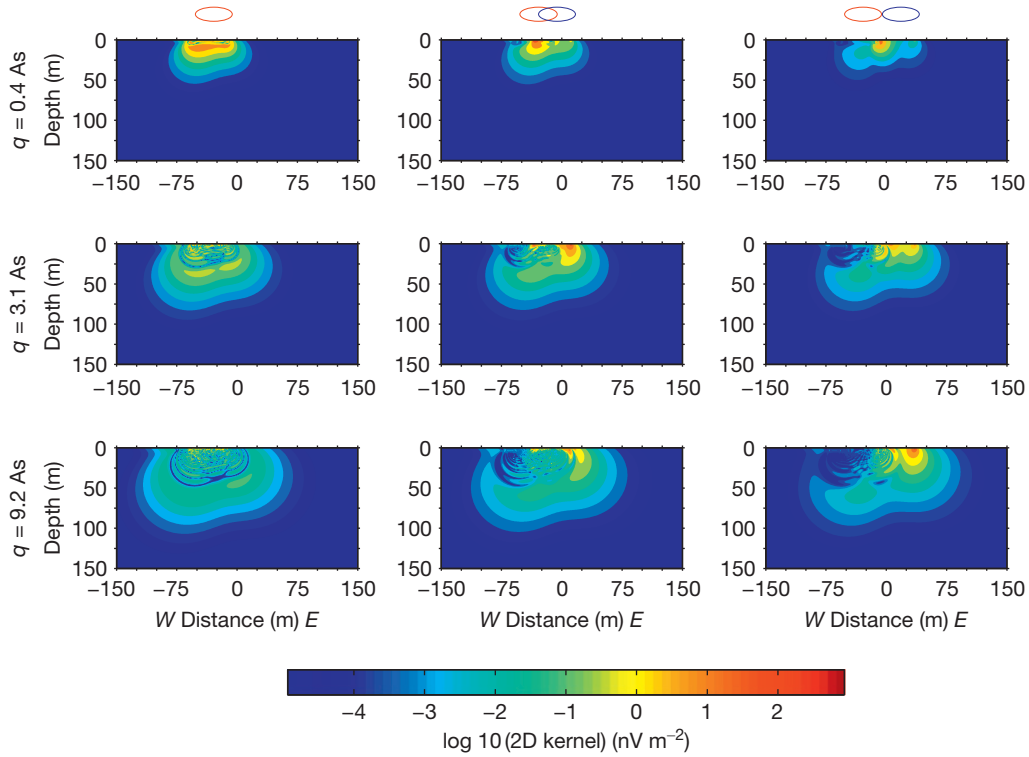


Figure 13 Two-dimensional sensitivity kernel for coincident loop and separated loop configuration. Kernels are calculated for the Earth's magnetic field intensity of 48,000 nT, inclination of 60°, and a half-space resistivity of 100 Ω m. Adapted from Hertrich M, Green AG, Braun M, and Yaramanci U (2009) High-resolution surface NMR tomography of shallow aquifers based on multi-offset measurements. Geophysics 74: G47–G59.

detected with depth. Increasing the pulse moment, the sensitivity reaches deeper regions but does not change in the lateral direction (Figure 13). Thus, the lateral resolution is limited to the spatial distance between the coincident loop measurements. In contrast, separated loops are sensitive to lateral change with increasing pulse moment. The larger the separation, the more distinct the sensitivity change in the lateral direction. However, with increasing separation, the S/N decreases because the detected signal becomes smaller. Therefore, only configurations up to edge-to-edge separation are considered. Unlike many other geophysical methods, transmitter and receiver are not interchangeable due to the inclination of the Earth's magnetic field. Thus, changing the transmitter and receiver adds new information, and consequently, a single measurement may consist of one transmitter and up to five simultaneous receivers. The profile consists of continuing to move this setup along the profile direction, repeating a measurement at a new transmitter position and inverting all together.

The state-of-the-art forward modeling includes arbitrary topography and three-dimensional distribution of conductivity (Lehmann-Horn et al., 2011), and initial attempts have also been made for 3-D inversion (Legchenko et al., 2011).

11.13.6 The NMR Link to Rock Properties

The aforementioned sections have provided a review on the fundamentals of NMR and their realization as laboratory, borehole, and field applications mainly from a physics perspective, that is, the perspective of measured quantities. As geophysicists, we are interested in rock properties that are necessary to describe geologic processes. Therefore, we will now focus on deriving properties such as porosity, permeability, and even tortuosity. Not all these rock properties can be addressed in the same kind in laboratory, borehole, and surface NMR. Especially, surface NMR is limited in many cases by the available experiments. Consequently, we note on how and under which conditions a property can be derived.

11.13.6.1 Sensitivity to Porosity

From eqn [4], the sensitivity of NMR to detect the number of hydrogen protons is easy to obtain, while the exclusive sensitivity to hydrogen protons is due the Larmor frequency (eqn [3]) being a function of the gyromagnetic ratio γ that depends on the actual nucleus. The initial amplitude, that is, the amplitude at $t=0$, of the measured relaxation signal is proportional to the water enclosed in a sample. In the case of laboratory or borehole NMR, a calibration measurement is utilized to transfer the amplitudes or measured voltages to water content. First, a sample containing a well-defined amount of water is measured. Second, the measured amplitudes of this calibration sample can be used to obtain the water content of an unknown sample.

In the case of surface NMR, such a simple calibration is not possible because the measured voltages depend on many variable factors, as shown in Section 11.13.5, such as loop size,

resistivity structure, and the Earth's magnetic field properties; however, most importantly, a depth-dependent calibration function is required. Therefore, a forward model to calculate the measured voltages for a subsurface water content distribution is used that allows the water content to be estimated from the measured data.

The use of measuring the water content is clear in the case of full saturation: it is a measure of porosity. However, partial saturation derived from water content, or relative changes of water content, can be of interest in characterizing samples of the near surface (Costabel and Yaramanci, 2011a,b).

11.13.6.2 Sensitivity to Permeability

11.13.6.2.1 A single pore

Starting with Seevers (1966), the relaxation time T_1 is described by a sum of two processes: the bulk relaxation T_B and a fast relaxation in a layer near a grain surface

$$\frac{1}{T_1}=\frac{1}{T_{1,B}}+r_s\frac{V_S}{V_B} \quad [23]$$

where r_s reflects this fast relaxation in the layer that is in contact with the grain surface, V_S is the volume of water in this layer, and V_B is the pore volume. It is the effect of the fast relaxation in this small layer enclosing the grains that gives rise to the estimate of the surface area to pore volume S/V of the pore using NMR relaxation times.

While Seevers did not rigorously treat the physics of this process, Brownstein and Tarr (1979) introduced a simple diffusion model calculating relaxation times T_1 for pores with basic shapes (planar, cylindrical, and spherical geometries). The model contains self-diffusion and surface relaxivity ϱ but neglects T_B. The material-specific rock property ϱ describes the interaction of the protons with the pore surface and corresponds to r_s introduced by Seevers (1966). The solution of the governing differential equation of the diffusion model is a sum of normal modes. For the main mode ($n=0$), which has the highest amplitude and longest relaxation time, the authors distinguished two limiting cases. For tube-shaped pores ($S/V=2/r_{tube}$ with a tube radius r_{tube}), the two cases are

- fast diffusion ($\varrho r_{tube}/D \ll 1$)

$$\frac{1}{T_1}=\varrho\frac{2}{r_{tube}} \quad [24]$$

- slow diffusion ($\varrho r_{tube}/D \gg 10$)

$$\frac{1}{T_1}=\frac{4D}{r_{tube}^2} \quad [25]$$

The solution for fast-diffusion conditions matches the formulation of Seevers (1966) without $T_{1,B}$. For slow diffusion, however, the relation of the relaxation time to the pore size changes and the increasing amplitude of the higher modes ($n>0$) add up to a significant amount of the total signal. This result leads to a multiexponential signal from a single pore.

Even though the aforementioned has been derived for T_1 relaxation times, T_2 is more common because the measurement is far less time-consuming. Note that T_2 relaxation times can yield identical pore size information (Kleinberg et al., 1993) if the influence of relaxation due to diffusion can be neglected.

But if NMR relaxation time is sensitive to a pore radius, what happens for a distribution of pore sizes?

11.13.6.2.2 Pore size distribution and multiexponential decay

In many cases, the observation of an NMR relaxation process in porous media reveals a multiexponential signal. Commonly, a smooth distribution of amplitudes A_i corresponding to a relaxation time distribution is estimated from the measured multiexponential signal (see **Section 11.13.3.5**). A_i is then transferred to water content by calibration.

There are two options to interpret the achieved smooth T_2 distribution: (i) The T_2 distribution represents a pore size distribution and the sum over an interval of A_i represents the partial water content of a pore size range and (ii) the T_2 distribution is affected by the higher modes of the NMR signal, that is, faster decaying components. Commonly, fast diffusion is assumed (e.g., Kleinberg et al., 1994) and validated for some sandstones and carbonate rocks (Latour et al., 1992). Thus, the distribution of relaxation times is interpreted as a pore radius distribution. **Figure 14** shows a comparison of scanning electron microscopy (SEM) images, computed tomography (CT) images, and NMR distribution of relaxation times. A straightforward approach for testing the assumption of fast diffusion is to utilize an independent measurement of the surface area, for example, gas absorption (Brunauer, Emmett, and Teller (BET) theory) measurements, and to calculate ϱ. However, there is discussion on the limitations for deriving ϱ to estimate the diffusion regime. It is beyond the scope of this chapter to provide further details; instead, we refer to Kleinberg (1996) or Kenyon (1997) discussing the effect of the technique used to measure the surface area, Hürlimann et al. (1994) describing the NMR intrinsic measurement of ϱ, and Sapoval et al. (1996) providing insight to the impact of more complex geometries as the simple ones used by Brownstein and Tarr (1979). Note that especially for near-surface sediments, comprehensive studies in this field have not yet performed.

11.13.6.2.3 The link to permeability

We start with a simple view on describing hydraulic conductivity K based on a model proposed by Kozeny (1927) and Carman (1938). The Kozeny–Carman model is based on a simple model of parallel tube-shaped pores with a radius r_{tube} allowing for the prediction of K of unconsolidated material. Even though Kozeny–Carman-type models are well accepted and used in hydrogeology, the application is, for instance, limited to laminar flow, not too wide grain-size distribution, noncompact particle shapes, and nonelectrochemical reactions (e.g., Carrier, 2003). The latter is why clayey material is not covered. Nevertheless, including a modification to estimate permeability k (e.g., after Pape et al., 2006), it is also used in rock physics. The original Kozeny–Carman equation reads

$$k_{KC}=\frac{1}{8T^2}\Phi r_{tube}^2 \quad [26]$$

where T is the tortuosity and Φ is the ratio of the void to total volume. Combining eqns [26] and [23] yields

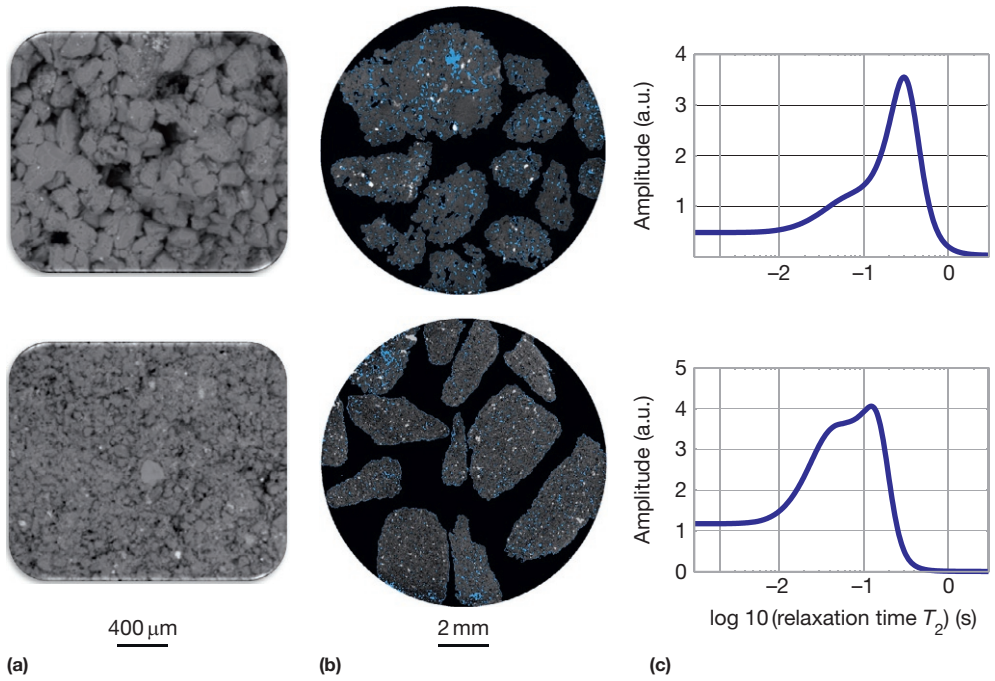


Figure 14 The sensitivity of NMR relaxation times to pore size. Measurements were conducted on cuttings taken from a deep geothermal drilling. (a) Scanning electron microscopy (SEM) images of two cuttings with different pore sizes; (b) computer tomography of several cuttings saturated with water (black represents air, blue represents water, and gray represents rock); (c) T_2 distribution estimated from NMR measurements at the cuttings in (b).

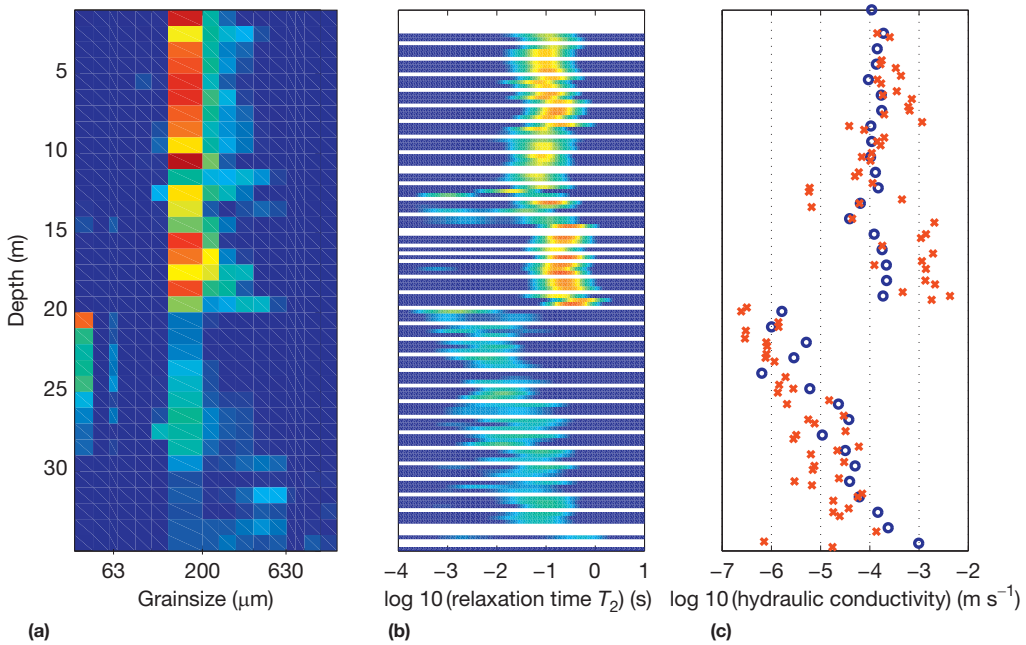


Figure 15 Comparison of sieving and NMR analysis from drilling material. (a) Grain-size distribution derived from sieving (cold/warm colors representing low/high weight percents), (b) NMR measured in the undisturbed sandy material within plastic tubes (cold/warm color represent low/high amplitudes); (c) hydraulic conductivities derived (a) based on grain sizes (red crosses) and (b) NMR (blue open circles).

$$k_S = C_S \Phi \left(\frac{T_{1,B} T_1}{T_{1,B} - T_1} \right)^2 \qquad [27]$$

where C_S comprises the matrix properties r_s and T. Note that T_1 might be replaced by T_2 resulting in a change of C_S. Figure 15 presents data obtained from sieving analysis and NMR measurements performed in undisturbed sandy material in plastic tubes extracted from drilling. The sieving analysis was performed according to common practice. The NMR relaxation time distribution matches the grain size very nicely. In

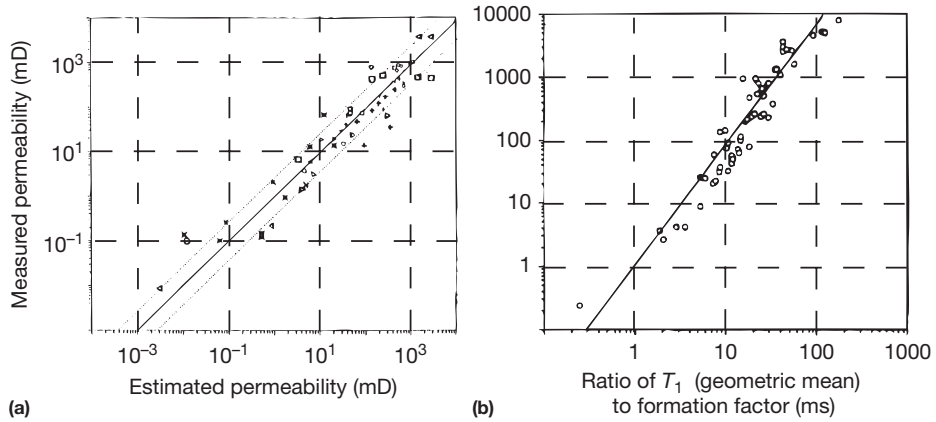


Figure 16 Measured permeability versus estimated permeability. (a) Permeability predicted as $k = 1.6 \times 10^{-9} \cdot T^{2.3}\phi^{4.3}$. Reproduced from Kenyon WE, Day P, Straley C, and Willemsen J (1988) A three-part study of NMR longitudinal relaxation properties of water-saturated sandstones: SPE Formation Evaluation 3: 622–636. (b) Permeability predicted as $k = 0.892 \cdot [T_1/F]^{1.77}$. Reproduced from LaTorraca et al. (1993).

hydrologic application, k is often replaced by the hydraulic conductivity K, which integrates the fluid properties dynamic viscosity and density. The NMR-based hydraulic conductivity is calculated using eqn [27] with adaption to K, and a single calibration factor is used to fit K derived after Carrier (2003) from the grain-size distribution.

In addition to this analytic derivation based on the simple model of tubes, an empirical equation to predict k from NMR measurements was presented by Kenyon et al. (1988). This equation is based on several datasets measured on sandstones and is commonly known as the Schlumberger–Doll Research (SDR) equation

$$k_{SDR} = C_{SDR}\Phi^a T_2^b \qquad [28]$$

with C_{SDR}, a and b determined from calibration. Kenyon et al. (1988) reported typical exponents of $a = 2$ and $b = 4$. Obviously, SDR and Seevers differ in the porosity exponent and by neglecting and accounting for bulk relaxation, respectively. Kenyon et al. (1988) reported that a porosity exponent of four is more sensitive to pore throats that control the permeability of rock, while Seevers (1966) was derived from the Kozeny–Carman type of flow originating from unconsolidated material. The latter may also exhibit large pores and therefore a significant impact of $T_{1,B}$. The impact of $T_{1,B}$ and slow diffusion on the estimation of K for coarse-grained unconsolidated material is reported, and an equation for improved prediction is presented recently by Dlugosch et al. (2013).

Equation [28] has been adapted in various ways but probably most interestingly linked to the formation factor known from Archie's law and then reads (LaTorraca et al., 1993)

$$k_{LT} = C_{LT}\frac{T_2^a}{F^b} \qquad [29]$$

with C_{LT} and a and b determined from calibration. Figure 16 shows two datasets comparing measured permeability with estimations based on NMR relaxation times.

As NMR signals may exhibit a multiexponential behavior originating from a distribution of pore sizes, how is this introduced into the given permeability estimators? An early approach of Timur (1969) accounted for three relaxation times and their respective weight by partial porosity. Meanwhile, it is common to calculate a logarithmic mean (Dunn et al., 2002) to calculate a representative $T_{1,2}$ that is then used as a representative constant. Unfortunately, there are several additional assumptions, such as pore coupling (Kleinberg, 1994), uniform surface relaxivity (Kenyon, 1997), and clay content (Kleinberg, 1996), that must be considered and may alter the prediction of k. However, the approach of Timur (1968) to handle bound water (as present in clay) is to derive the irreducible water saturation S_{iw} (also used in terms of bulk volume of irreducible saturation (BVI) $= \phi S_{iw}$) from relaxation times below a certain threshold, for instance, 33 ms (Straley et al., 1997), and estimate permeability according to

$$k_{TC} = C_{TC}\left(\frac{\Phi^{4.4}}{S_{iw}^2}\right) \qquad [30]$$

Equation [30] is usually referred to as Timur–Coates equation and C_{TC} is a calibration constant.

While both T_1 and T_2 are available from laboratory and borehole measurements and therefore the given equations can be applied, surface NMR is mostly restricted to FID measurements. There are few papers on retrieving permeability from FID measurements at the field scale. For an overview, see, for instance, Mohnke and Yaramanci (2008), Plata and Rubio (2008), and references therein. But T_2^* is sensitive to inhomogeneities in the magnetic field, that is, magnetic gradients at the pore scale reduce T_2^* (Grunewald and Knight, 2011) and therefore significantly affect the prediction of k. Thus, obtaining relaxation times other than T_2^* is of high interest for surface NMR. Latest developments, therefore, focused on developing reliable T_1 and T_2 measurements and inversion schemes and promise to improve the prediction of k from surface NMR.

11.13.6.3 Unsaturated Hydraulic Conductivity: Exploring the Vadose Zone

Understanding hydraulic conductivity under partially saturated conditions is of great importance for water transportation characteristics in the vadose zone. Costabel and Yaramanci (2011a) presented an approach to estimate relative hydraulic conductivity connecting NMR and the semiempirical model of Brooks and Corey (1964). Based on the characterization of the

water retention curve by the Brook–Correy model, relative hydraulic conductivity as ratio of unsaturated (K_U) to saturated hydraulic conductivity (K_S) is given as

$$K_r = \frac{K_U}{K_S} = S_E^{a+(2/\lambda)} \quad [31]$$

with S_E being the effective water saturation, λ the pore size-distribution index, and a an empirical parameter related to the tortuosity of the medium. Costabel and Yaramanci (2011a) showed that K_r can be estimated by means of NMR using

$$K_r = S_{E,NMR}^{a} \cdot T_{2,rel}^{2} \quad [32]$$

where $S_{E,NMR}$ is the effective saturation derived from NMR and $T_{2,rel}$ the relaxation time at the corresponding saturation degree. The authors validate eqn [32] by calculating relative hydraulic conductivities using Hydrus-1D (Simunek et al., 1999) based on measuring the outflow of a drainage experiment and comparison to the NMR results (Figure 17).

11.13.6.4 Petrophysical Characterization Based on Diffusion Measurements

11.13.6.4.1 Surface-to-volume ratio

As mentioned in Section 11.13.3.6, there are a second set of experiments in addition to relaxometry that allow for exploring pore space properties. These experiments are used to determine the self-diffusion coefficient D of the pore fluid. Mitra et al. (1992) and Hürlimann et al. (1994) demonstrated that the ratio of the surface area to pore volume (S/V), among other parameters, can be estimated using measurements of the time-dependent effective diffusion coefficient $D(t)$, that is, to obtain the mean square displacement of a proton. In a restricted environment, that is, the pore space, the effective diffusion depends on the duration the proton is allowed to move. Initially, $D(t)$ equals the diffusion coefficient of free water D_0 because the proton is free to move. At short times, the proton experiences the presence of the pore (Figure 18), and the ratio

$$\frac{D(t)}{D_0} = 1 - \frac{4}{9\pi}\frac{S}{V}\sqrt{D_0 t} + \Theta(D_0 t) \quad [33]$$

is a measure for S/V, with $\Theta(D_0 t)$ representing higher-order terms. These results are of great interest because k estimates from relaxation times always depend on several assumptions and knowledge/calibration of surface relaxivity. Thus, a second independent measurement (that has its own limitations) is advantageous. Measuring restricted diffusion not only allows S/V to be estimated but also provides information on tortuosity (Latour et al., 1995; Pape et al., 2006), another parameter that controls k. Tortuosity is derived from the late time behavior of $D(t)$:

$$\frac{D(t \to \infty)}{D_0} = \frac{1}{T} \quad [34]$$

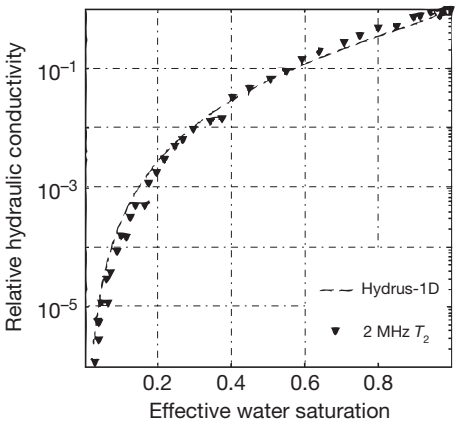


Figure 17 Relative hydraulic conductivity as a function of effective water saturation for a medium sand sample. Hydrus-1D results (dashed line) are obtained by means of outflow measurements from a drainage experiment. NMR results (rectangles) are based on eqn [32] and conducted at 2 MHz. Adapted from Costabel S and Yaramanci U (2011) Relative hydraulic conductivity in the vadose zone from magnetic resonance sounding—Brooks–Corey parameterization of the capillary fringe. Geophysics 76: G61–G71.

11.13.6.4.2 Pore fluid characterization

Further exploration of diffusion properties gives rise to pore fluid characterization (Freedman and Heaton, 2004; Winkler et al., 2005) or even the wettability of fluids in the pore (Freedman et al., 2002; Zhang et al., 2000). Especially, multi-dimensional (D-T2) experiments allow for exploring petrophysical properties in great detail.

An example of pore fluid identification based on D-T2 maps as presented by Flaum et al. (2005) is given in Figure 19. It shows that, if only relaxation times are measured, identification of the pore fluid is not possible since the relaxation times of water and oil are too close in that case. If diffusion coefficients alone are measured, there is ambiguity since restricted diffusion may be present. It is the combined multidimensional experiment that allows for a unique characterization. There is even more to gain. If the measured diffusion coefficients are above the bulk diffusivity, then internal gradients are present. If the measured diffusion coefficients are below, then restricted

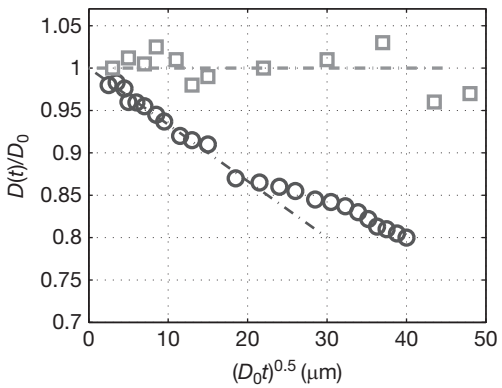


Figure 18 Restricted diffusion for a monosized pack of glass spheres (open circles) and bulk fluid (open squares). Lines represent a linear fit in the interval [0.85 1] representative of the short-time behavior of $D(t)$. The slope indicates the S/V of the glass bead system. Adapted from Vogt C, Galvosas P, Klitzsch N, and Stallmach F (2002) Self-diffusion studies of pore fluids in unconsolidated sediments by PFG NMR. Journal of Applied Geophysics 50: 455–467.

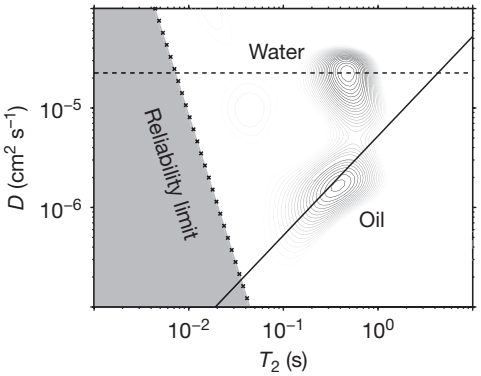


Figure 19 D-T2 map of a sandstone saturated with water and oil. The diffusion coefficient of bulk water is given as dashed line and the solid diagonal line with positive slope indicates the correlation between diffusivity and T_2 for hydrocarbon mixtures. The second diagonal line with negative slope indicates the limits of interpretation (shaded area should not be used for interpretation). Adapted from Flaum M, Chen J, and Hirasaki GJ (2005) NMR diffusion editing for D-T2 maps: Application to recognition of wettability change. Petrophysics 46: 113–123.

diffusion occurs. Furthermore, if restricted diffusion occurs at long relaxation time, then surface relaxivity is low.

Wettability can be determined by comparing the relaxation time of the bulk fluid with the response measured from a saturated rock sample. Since relaxation time is essentially influenced by the interaction at the pore surface, a nonwetting fluid will show the same relaxation time response measured either as bulk fluid or within a pore.

Another branch of multidimensional experiments is T1–T2 correlation maps (Song et al., 2002) to distinguish the porosity of large pores from interparticle porosities, among other types of information (Schoenfelder et al., 2008).

Obviously, these rock properties are of special interest for oil and gas exploration. Even though pulse field gradients are not available in borehole NMR tools, the constant gradient is used to explore diffusion effects. With respect to surface NMR, these types of experiments are impossible because neither gradients in $\mathbf{B}_0$ nor appropriate CPMG sequences are yet available.

11.13.7 Case Histories

The following section gives real-world examples covering different aspects of NMR. The first case history shows how laboratory NMR provides complementary information to a suite of borehole logs and laboratory measurements in order to complete a site characterization in the context of geothermal exploration. The second example shows how borehole NMR logging in a groundwater monitoring well compares to a traditionally made geologic description that is based on analyzing core material. The third case compares borehole NMR and surface NMR and shows the ability of using borehole NMR and surface NMR jointly to extend 1-D borehole-derived information in lateral direction. The fourth case demonstrates the joint use of surface NMR and resistivity methods to solve the ambiguities of resistivity methods to distinguish clayey material from saltwater intrusion. Finally, the fifth example shows the ability of surface NMR to distinguish frozen from unfrozen sediments.

11.13.7.1 Hard Rock Reservoir

11.13.7.1.1 Objective

The aim of the study published by Hübner et al. (2013) is to characterize possible layers for geothermal use in terms of their physical and sedimentary petrographic properties using borehole logging and laboratory measurements. As a demonstration project for geothermal energy in regions with low-permeability sedimentary rocks and to evaluate the capabilities of single wells for geothermal purposes, a single deep well was drilled in northern Germany. It reached its final depth at 3901 m below ground level. The target for geothermal exploration is the sandstones of the mid-Lower Triassic at depths between 2300 and 2800 m. The sandstone formation (German Wealden) at depths between 1145 and 1335 m is intended to serve as buffer storage for the cyclically injected and produced water. A 14.5 m core (1208.2–1222.7 m) and nine sidewall cores from several meter thick sandstone strata were available for laboratory analyses, and downhole logging began at 1175 m for logging analysis. In particular, laboratory NMR was carried out to determine the volume of nonextractable water (BVI) that is obviously an important petrophysical parameter in the context of geothermal energy.

11.13.7.1.2 Methodology and results

Logging was performed for the following parameters: natural radioactivity (GR), well diameter (CAL), seismic velocity (Vp), density (DEN), neutron porosity (NPOR), and photoelectric factor (PEF). Cluster analysis was used to determine the lithology based on the downhole logs, which resulted in a separation of five clusters (**Figure 20**). Laboratory analysis was performed to obtain the porosity using the buoyancy method, including measuring the mass of the dry sample, the saturated sample, and the sample in water. Permeability was determined using a special device for high-resolution gas-permeability measurements (Westphal Präzisionstechnik GmbH & Co. KG, Celle, Germany) in up to three spatial directions depending on sample shape. Information on the pore size distribution and porosity was obtained from NMR relaxometry (MARAN Ultra 5, Oxford Instruments, Oxfordshire, UK). The volume of nonextractable water was calculated by summing up the T_2 intensities below 33 ms (Miller et al., 1990).

Clusters 1 and 2 are sandstones. The sedimentary petrographic analyses reveal predominantly quartzitic sandstones with similar components and considerable variation in terms of cementing, matrix percentage, and grading. The average porosity of the sandstone plugs is 14%. The plugs taken from the drill core display axial permeability values of 50 mD on average. Some radial core plugs and many sidewall cores are even more permeable (**Figure 20**). Most of the T_2 distributions exhibit dominant fractions with an intensity of over 50 ms at times, indicating the presence of large pores. Both narrow and broad distributions were observed, mirroring the variety of pore size distributions of the sandstones. Most samples with a short average T_2 have porosities below 15% and a BVI of over 50 %, that is, more than half of the water is stored in small pores. For half of the studied sandstone samples, the BVI was determined below 50% and even less than 20% for some sidewall cores. Sandstone 1 is moderately to well-graded and

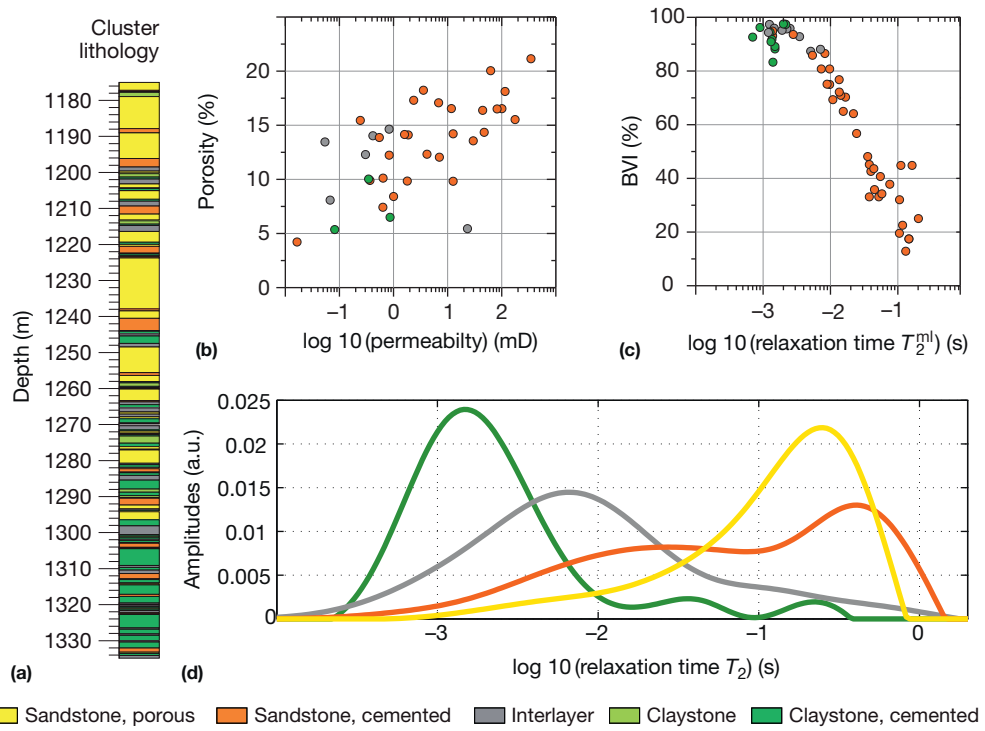


Figure 20 Characterization of a target layer for geothermal energy from borehole logging and laboratory measurements. (a) Lithology based on a cluster analysis of logging data. (b) and (c) Cross plots of laboratory-derived permeability versus porosity and bulk volume of irreducible (BVI) saturation versus mean log. T_2^{ml}. (d) NMR T_2 distributions to characterize pore sizes. Figure compiled from Hübner et al. (2013).

medium- to fine-grained, while sandstone 2 is moderately to poorly graded and medium- to fine-grained.

Clusters 4 and 5 are clay stones. Clay stone 1 is characterized by lower PEF values and higher DEN values compared with clay stone 2. In contrast to clay stone 1, clay stone 2 has an iron oxide/hydroxide cement in places, which originates from the diagenetic formation of iron-rich minerals, such as a biotite. It was not possible to distinguish the two clay stones based on the laboratory measurements. According to the T_2 distribution, the clay stone samples have small pore sizes. Almost the entire amount of water (up to 98%) is nonextractable. With regard to permeability measurements, most of the values were disregarded because the results only reflected the permeability along fractures that are unexpected for the in situ material and most likely developed during and after the drilling process. The remaining values are below 1 mD.

The finely layered interbedding (cm range) of sandstone and clay stone/siltstone cannot be differentiated in more detail due to the limited vertical resolution of the downhole logging (sampling rate: 15 cm).

11.13.7.1.3 Discussion

An important parameter for reservoir characterization is the transmissibility, that is, the product of radial permeability and thickness. Transmissibility is generally determined from hydraulic tests; however, for an initial estimation, this parameter may be calculated using the data acquired in the laboratory, such as NMR data on pore size distribution and BVI, and lithology-related thicknesses determined with the aid of cluster analyses. The deciding factor affecting the overall transmissibility is the thickness of the permeable layer, whereas the less permeable layers are only of minor importance because horizontal flow is examined. According to this preliminary estimate, the formation exhibits the needed transmissibility for an efficient production at the location of the well and should be envisaged as a potential buffer storage for the cyclically injected and produced water. Nevertheless, in situ tests within the formation are imperative to validate these statements.

11.13.7.2 Sedimentary Aquifer

11.13.7.2.1 Objective and methodology

The aim of this study is to evaluate a slim-borehole NMR tool (Javelin, Vista Clara) capable of performing NMR logging in wells as slim as 2 in. in diameter. The test site is located in northern Germany and mainly consists of unconsolidated glacial sediments. The site has been characterized using other geophysical methods (e.g., resistivity and georadar) (see Mueller-Petke and Yaramanci, 2010) indicating a primary aquifer of medium to coarse sand down to about 11 m separated from a second confined aquifer by a glacial till layer. Additionally, a thin peat layer is identified at about 5 m depth. A 2 in. monitoring well, which is common in many near-surface hydrologic applications, was drilled, and the core material was analyzed. In addition to NMR logging, a slim gamma-ray tool was used to identify the fine-grained material.

11.13.7.2.2 Results and discussion

Figure 21 shows the lithology derived from drilling material, the gamma-ray log, and T_2 distribution from NMR logging. Both gamma-ray and NMR clearly identified the aquiclude structure (till layer interbedded with sand) by increased gamma activity

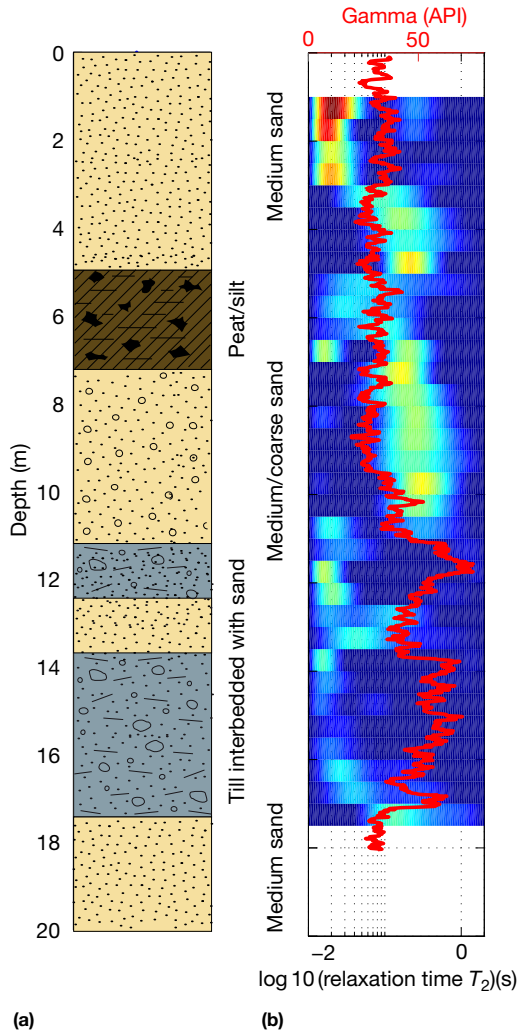


Figure 21 Characterization of a sedimentary aquifer from borehole logging in a 2-in. monitoring well. (a) Lithology derived from drilling and additional geophysical measurements; (b) NMR and gamma-ray logging (red line). In (b), the color code refers to the water content with warm/cold colors indicating high/low values.

and reduced T_2 relaxation time. These results correlate well with the known lithology. Note that the current 2 ms echo spacing of the Javelin does not allow bound water to be detected with a high level of accuracy; however, depending on the S/N, bound water can at least be detected. However, NMR allows additional information to be obtained. The aquifer can be identified and characterized to a depth of 11 m. The upper part of the aquifer (down to 5 m) is dominated by medium sand, and the second part below the peat layer contains a significant amount of coarse sand, causing slightly higher relaxation times. While the peat layer is invisible for gamma-ray logging, T_2 relaxation times are reduced, as up to 50% of the water in peat is immobile (Culligan and Sinfield, 2001).

11.13.7.3 Groundwater Storage

11.13.7.3.1 Objective

Müller-Petke et al. (2011b) presented a study carried out at a groundwater test site for aquifer storage and recovery in the desert of Abu Dhabi, where borehole NMR logging and surface NMR along a profile were conducted in order to (i) extend the hydrogeologic knowledge of the area, (ii) demonstrate the comparability and synergy of both NMR techniques, and (iii) evaluate the potential of surface NMR-derived results in comparison with that of borehole NMR-derived results.

11.13.7.3.2 Methodology and results

The test site is located in the emirate of Abu Dhabi in the United Arab Emirates (UAE), near the border to Oman approximately 60 km north of the city of Al Ain. The geology at the site is mainly dominated by a tightly folded and thrust-faulted regime striking NNE–SSW originating in the Late Eocene–Miocene collision of the northeastward-moving Arabian Plate and the Eurasian Plate. The near-surface sediments can be divided into Quaternary unconsolidated aeolian and fluvial sands on top of a clay stone with interbedded dolomitic marl, limestone, and siltstone. The main target for the aquifer storage project is the saturated aquifer (≈40–60 m below surface) that is part of the unconsolidated sediments.

A 3-D TEM survey carried out in advance covers a larger area of the test site. The positions and results of selected TEM soundings along a profile are shown in Figure 22(a) and 22(b). The unsaturated aquifer is potentially characterized by resistivities larger than 100 Ω m, while the saturated aquifer is indicated by a decrease of resistivity to about ±30 Ω m. The clayey basement then shows resistivity lower than 5 Ω m. A clear identification of the aquifer extension and thickness is, however, challenging. A discontinuity in the resistivity structure of the clayey basement between SO71 and SO73 is visible. According to the complete 3-D TEM survey, an anomaly is striking NNW–SSE and potentially separates the aquifer storage structure. The NMR survey consists of a borehole NMR log using the CMR tool (Schlumberger, Houston, USA) at W1 and 11 surface NMR soundings using the NUMIS Plus (Iris Instruments, Orleans, France). The profile layout was chosen to cross the expected discontinuity in the clayey basement. An overview of all positions is given in Figure 22(a).

Figure 22(c) and 22(d) displays the borehole NMR logging for W1 together with surface NMR at P1. While borehole NMR is presented in terms of T_2 relaxation time distribution and water content integrated over different T_2 intervals, surface NMR gives a mean T_2^* relaxation time (not shown here) and the total detected water content. Based on logging, the aquifer bottom is clearly identified by a decrease in relaxation times of about two decades within a transient zone from about 215 m down to 205 m. The aquifer top is characterized by an increase in water content.

In order to compare borehole NMR water content with surface NMR, two intervals to integrate relaxation times are selected, one of which displays water content related to relaxation times below and the other above 60 ms. This cutoff was selected since the dead time of 40 ms and a 40 ms pulse length do not allow for the detection of relaxation faster than 60 ms. Taking this into account, the water content estimated by surface NMR agrees well with the ones derived from the logging, while the vertical resolution of surface NMR is much lower and does not permit a detailed interpretation. Nevertheless, considering the comparison of surface NMR and borehole NMR, the measurements along the profile allow for extending this 1-D borehole information to image the lateral extent of the

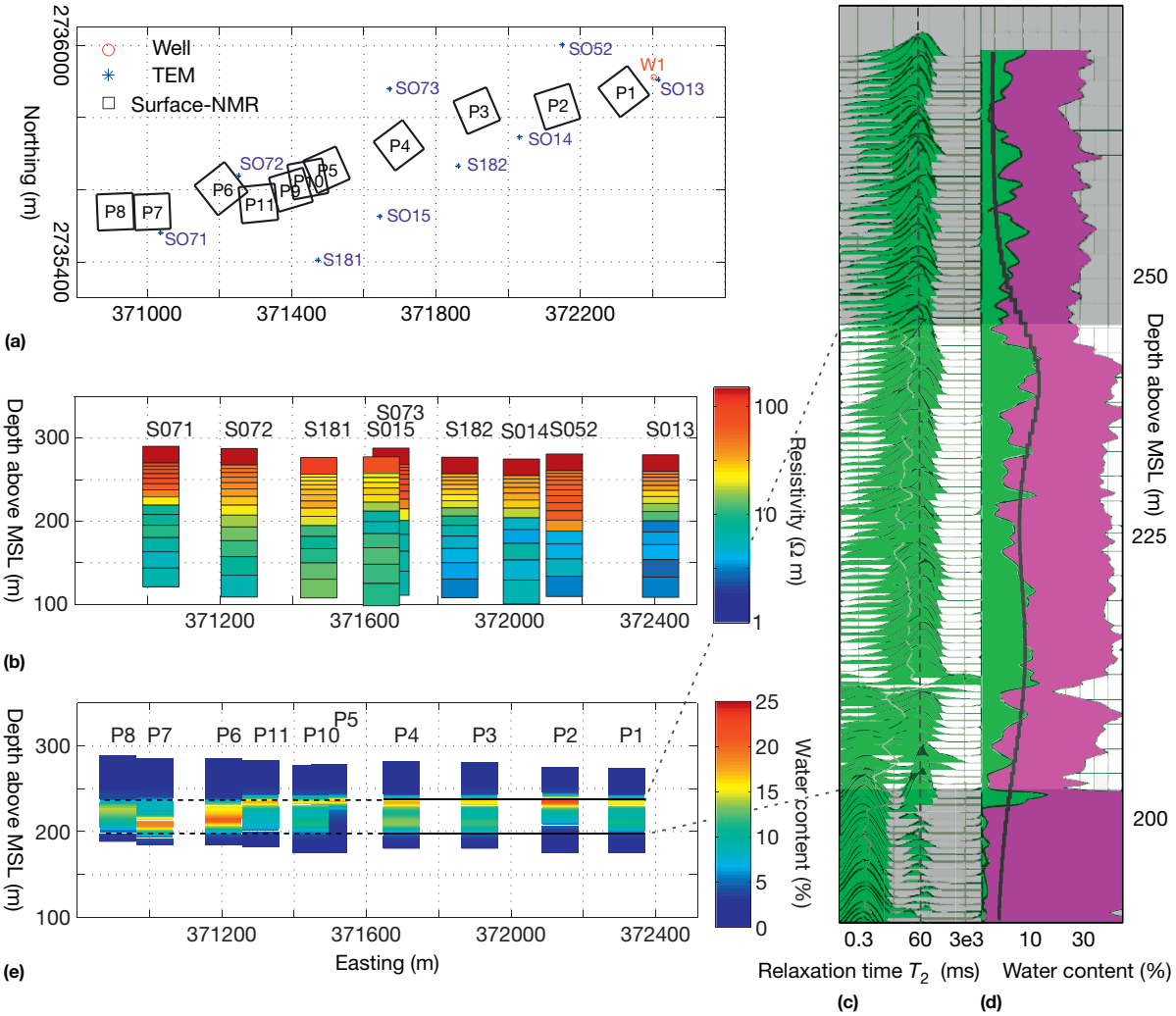


Figure 22 Characterization of a groundwater storage test site. (a) Positions of TEM, surface NMR and wells. (b) Resistivity along the profile as derived from the TEM. (c, d) NMR logging at W1 compared to surface NMR at P1 with (c) T_2 distributions with depth and (d) water content with depth. In (d), the color code refers to integration over intervals of the T_2 relaxation time distribution to derive the water content corresponding to this interval. The respective intervals are magenta (0.3–60) ms and green (60–3000) ms. The solid lines show the water content as from surface NMR. The dashed line indicates the 60 ms used to separate the intervals. (e) Water content along the profile as derived from surface NMR. Figure compiled from Müller-Petke et al. (2011b).

aquifer using surface NMR (Figure 22(e)). Similar to the TEM results, an anomalous area is potentially indicated by changed lower aquifer boundaries between P5 and P11. There is an indication, especially at P5, that the aquifer might be separated, but due to the fast-decaying signals (80 ms in average) and low S/N of the surface NMR data, there is unfortunately high uncertainty associated with these findings. Nevertheless, both the aquifer top and the aquifer bottom are clearly imaged and provide valuable information for groundwater modeling, that is, NMR provides required hydrologic information such as aquifer thickness, water content/porosity, and potentially hydraulic conductivity.

According to the interpretation of the logging data, we emphasize that apparently, the water content in the partially saturated zone on top of the aquifer is overestimated. Assuming that the sensitive volume, that is, the depth of investigation, of the tool is not completely within the undisturbed formation but partly measuring borehole fluid infiltrating the formation might be a plausible explanation to this. Furthermore, note that, even though site-specific, a cutoff T_2 of 33 ms is typically used as a first guess to distinguish nonextractable (or bound) water from extractable (or free) water. As a consequence in the given example, most of the water content in the aquifer belongs to free water.

11.13.7.4 Saltwater Intrusion

11.13.7.4.1 Objective

Climate change affects groundwater dynamics, particularly at the coast and on barrier islands, where the interaction between freshwater and saltwater is important for the water supply. Airborne electromagnetic measurements are particularly important because they provide a 3-D distribution of electrical conductivity, which is a key parameter in hydrogeophysical

investigations. However, there are two main shortcomings of resistivity: (i) We cannot clearly differentiate between clay content and fluid salinity and (ii) there is no sufficiently reliable relation to hydraulic conductivity.

11.13.7.4.2 Methodology and results

Günther and Müller-Petke (2012) published an approach to overcome these shortcomings by joint inversion of resistivity and NMR. Both datasets are jointly inverted for resistivity, water content, and relaxation time using a joint inversion scheme. Coupling is accomplished via common layer thicknesses. The two sounding positions (at each position, NMR and DC sounding were conducted) are located in the center (OD33) and at the boundary of the freshwater lens (SKD). The latter is near the North Sea. In this flat area, storm surges regularly add saltwater from the top. Because the lithology consists of fine to medium sand layer with interbedded silt/clay layer, the task of distinguishing the clay from saltwater-saturated sand is of great importance in understanding groundwater dynamics.

The joint inversion results for both sounding positions are displayed in Figure 23. The error bars are calculated based on the parameter variations (thickness, water content, resistivity, and relaxation time of each layer) until the data cannot be sufficiently satisfied within the data error. At both positions, the water content is ~35% and almost constant with depth. Resistivity and relaxation times together, however, permit a clear picture of the subsurface that cannot be derived from each method alone. While a decrease of resistivity without a change in relaxation times indicates a change of pore fluid, decreasing relaxation time and low resistivities indicate a fine-grained material. Thus, the aquifers at OD33 are not affected by saltwater. The decrease in resistivity at 25 m is due to the increase of fine-grained material. Saltwater-saturated silt is present below 55 m. In contrast to OD33, at SDK, a saltwater-saturated aquifer is identified below 30 m. It further shows that the saltwater originating from the flooding events does not immediately reach the freshwater lens (low resistivities but high relaxation times at about 5 m depth) and comparing resistivities and relaxation times between 10 and 30 m between OD33 and SDK, the aquifer is rather brackish.

11.13.7.4.3 Discussion

The joint inversion not only significantly improves the reliability of the results (as the authors show in their original publication) but also helps to distinguish fluid conductivity from lithology. We would like to emphasize that short T_2^* relaxation times with no change in resistivity indicate faster relaxation due to magnetic inhomogeneities, as the authors demonstrated for another field example. It will be of great importance to conduct T_1 experiments based on the newly developed pcPSR (Walbrecker et al., 2011b) to obtain relaxation times that are less affected by magnetic gradients.

11.13.7.5 Permafrost

11.13.7.5.1 Objective

A talik is a layer or body of unfrozen ground that occurs in permafrost due to an anomaly in thermal, hydrologic, or hydrochemical conditions. These unfrozen sediments or talik may extend many tens of meters. Thermokarst lakes, formed due to thawing of permafrost followed by subsidence, can emit considerable volumes of methane. A talik can be an isolated body in the subsurface, a closed body below a thermokarst

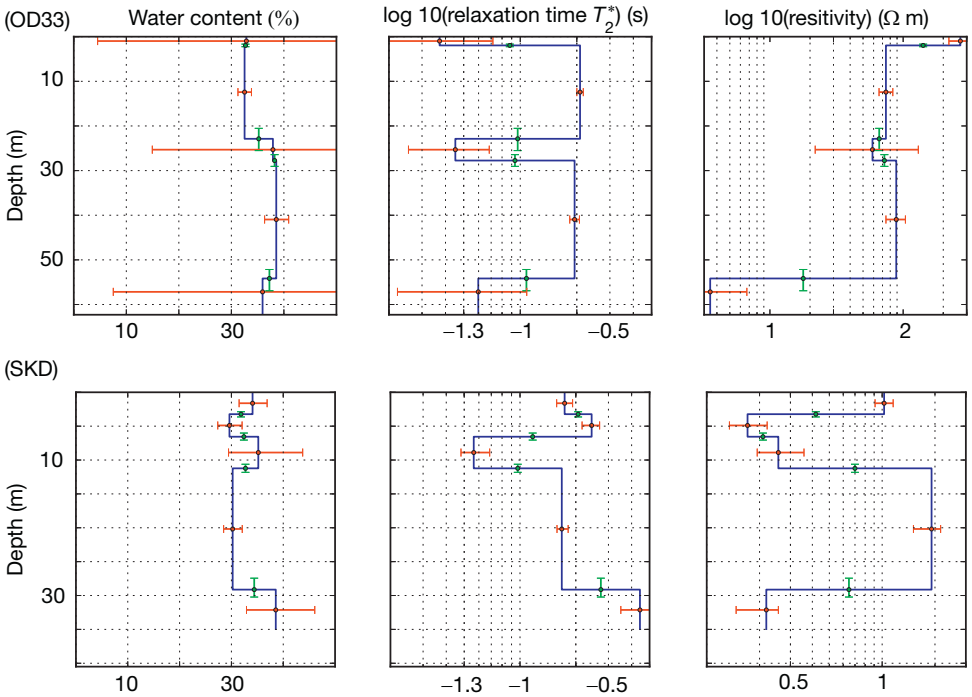


Figure 23 Identification of saltwater intrusion using joint inversion of MRS and VES. Soundings are located in the center (a – OD33) and at the boundary of a freshwater lens (b – SKD). Figure compiled from Günther and Müller-Petke (2012).

lake, or an open body with connection to the regional groundwater. Thus, information about the geometry of the unfrozen body is important to understand the regional water dynamics.

11.13.7.5.2 Methodology and results

Parsekian et al. (2013) studied this phenomenon in Fairbanks, Alaska, using surface NMR. Because surface NMR cannot detect frozen water due to its fast relaxation, the method offers a unique sensitivity to the unfrozen sediments. Figure 24 displays the modeling of plausible scenarios to demonstrate the capability of the method to localize a talik even if a thermokarst lake on top is present. Figure 25 shows the results of three soundings in Fairbanks, Alaska, using a layered inversion. The first sounding (Bonanza Creek) was performed at a terrestrial permafrost site with the nearby borehole that indicates that the permafrost is at least 24 m deep. The second (Caribou Lake) and third (Ace Lake) soundings were conducted on top of a frozen thermokarst lake with an expected talik below, while the current size and depth of Ace Lake suggest considerable thaw subsidence and a relatively long period of talik expansion in the ice-rich permafrost of this area. Based on the NMR results, the lake size and depth, and the permafrost thickness, the authors conclude that Ace Lake has either a closed talik that extends deeper than the maximum depth of their measurement (22 m) or an open talik, potentially connecting the lake to the subpermafrost water. The data from Caribou Lake best describe a system in which a talik extends to a depth of 4–6 m below the shallow lake and is underlain by permafrost. At Bonanza Creek, the results reveal a layer of elevated water content at a 40 m depth that is bounded above and below by low-water content layers.

11.13.7.5.3 Discussion

The study demonstrates the ability of surface NMR to detect unfrozen sediments below thermokarst lakes of different depths, estimate the depth of taliks, and detect unfrozen sediments below permafrost. Surface NMR is one of the preferable geophysical investigation methods for detecting unfrozen sediments in permafrost regions because of its unambiguous, direct sensitivity to liquid water.

11.13.8 Summary

This chapter gives a basic description on the fundamental physics of NMR and its implementation as laboratory, borehole, and field technique in the context of geosciences. Mostly, a qualitative description and analogies are preferred to invoke a more intuitive understanding rather than following a detailed mathematical or even quantum mechanical derivation that can be found elsewhere. Thus, we focus on basic experiments and the governing concepts to provide access to the technique, to understand the effects of the main controlling parameters, and to hopefully motivate the well-disposed reader to progress with further reading.

Special attention is paid to clearly distinguish the governing concepts of laboratory, borehole, and surface NMR in their respective kind of conducting an NMR experiment, while less attention is paid on the differences in scales that are involved in an NMR experiment. We leave this field of rock physics open, but note that a technique that finds application to discover properties from micrometer scale to catchment scale is potentially suitable to target questions related to scale properties.

A couple of field examples are presented introducing some aspects of NMR to the reader, such as permeability estimation or imaging of water content. But we like to emphasize that there are a lot of applications that are not explained, such as imaging of root growth in soils or fluid flow in rock cores. It is one of the fascinating aspects of NMR that the range of applications appears to be infinite. Besides this fascination

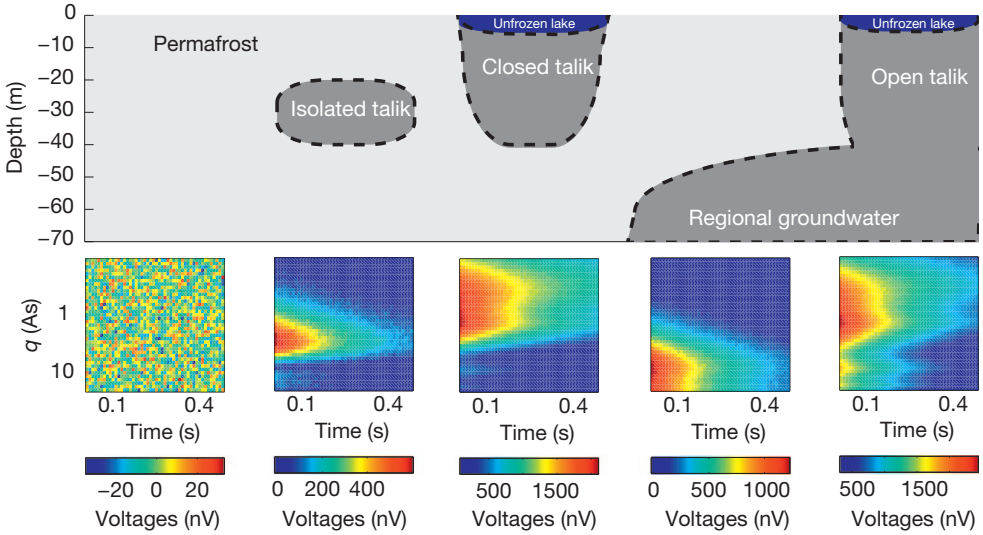


Figure 24 Modeling of a plausible talik scenario to demonstrate the sensitivity of the applied method. The modeling was performed using a 100 m diameter circular loop and a 48,000 nT Earth's field at a 60° inclination. The data were contaminated with 10 nV Gaussian distributed noise. For permafrost, we assume 30% water content and $T_2^* = 1e-5$ s relaxation time, for the lake 100% water content and $T_2^* = 1$ s and talik 30% water content and $T_2^* = 0.3$ s.

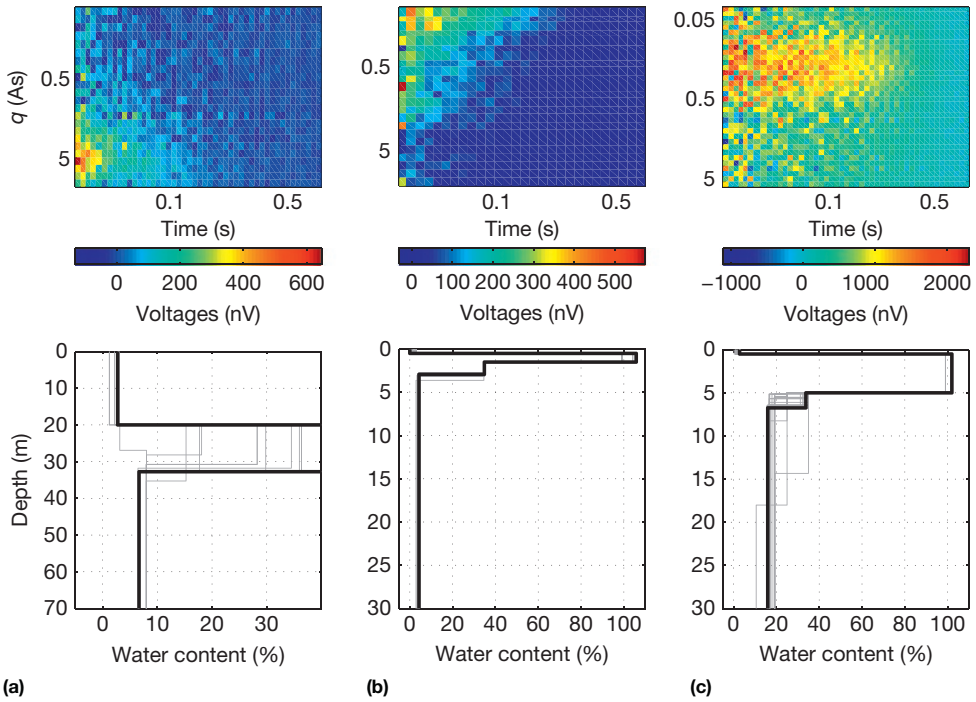


Figure 25 Field data of thermokarst anomalies at Bonanza creek (a, 100 m square loop), Caribou Lake (b, 50 m diameter figure eight), and Ace Lake (c, 50 m diameter figure eight) near Fairbanks, Alaska. The black lines are the best-fitting models, while the gray lines represent equivalent results and therefore indicate the model uncertainty. Redrawn from Parsekian AD, Grosse G, Walbrecker JO, et al. (2013) Detecting unfrozen sediments below thermokarst lakes with surface nuclear magnetic resonance. Geophysical Research Letters. http://dx.doi.org/0.1002/grl.50137.

concerning the unique possibilities of NMR, we finally like to point out that any technique comes along with limitations and assumptions and should always be used complementary with other methods in best practice.

Acknowledgments

The authors would like to thank the NMR and hydrogeophysics group at the Leibniz Institute for Applied Geophysics and Technical University of Berlin for many years of fruitful discussion, laboratory and field work, and the data provided for this work. We would particularly like to thank the hydrogeophysics group of R. Knight at Stanford University for providing the Alaska dataset. The visit of MMP at Stanford University was supported by Blaustein Visiting grant to R. Knight. Finally, we like to acknowledge the work of three anonymous reviewers and Lee Slater that significantly helped to streamline the chapter.

References

Behroozmand AA, Auken E, Fiandaca G, Christiansen AV, and Christensen NB (2012) Efficient full decay inversion of MRS data with a stretched-exponential approximation of the distribution. *Geophysical Journal International* 190: 900–912.

Bloch F (1946) Nuclear induction. *Physical Review* 70: 460–474.

Bloembergen N, Purcell EM, and Pound RV (1948) Relaxation effects in nuclear magnetic resonance absorption. *Physical Review* 73: 679–712.

Braun M, Hertrich M, and Yaramanci U (2005) Study on complex inversion of magnetic resonance sounding signals. *Near Surface Geophysics* 3: 155–163.

Braun M and Yaramanci U (2008) Inversion of resistivity in magnetic resonance sounding. *Journal of Applied Geophysics* 66: 151–164.

Brooks RH and Corey AT (1964) *Hydraulic Properties of Porous Media: Hydrology Papers*. Fort Collins, CO: University Colorado State.

Brown R and Gamson B (1960) Nuclear magnetism logging. *Journal of Petroleum Technology* 219: 199–201.

Brownstein KR and Tarr CE (1979) Importance of classical diffusion in NMR studies of water in biological cells. *Physical Review A* 19: 2446–2453.

Callaghan P (2007) *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford: Oxford University Press.

Carman PC (1938) The determination of the specific surface of powders. *Journal of the Society of Chemical Industrialists*. 57: 225–234.

Carr HY and Purcell EM (1954) Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review* 94: 630–638.

Carrier WD (2003) Goodbye, Hazen; Hello, Kozeny–Carman. *Journal of Geotechnical and Geoenvironmental Engineering* 129: 1054–1056.

Cooper RK and Jackson JA (1980) Remote (inside-out) NMR. I. Remote production of region of homogeneous magnetic field. *Journal of Magnetic Resonance (1969)* 41: 400–405.

Costabel S and Yaramanci U (2011a) Relative hydraulic conductivity and effective saturation from Earths field nuclear magnetic resonance a method for assessing the vadose zone. *Near Surface Geophysics* 9: 155–167.

Costabel S and Yaramanci U (2011b) Relative hydraulic conductivity in the vadose zone from magnetic resonance sounding—Brooks–Corey parameterization of the capillary fringe. *Geophysics* 76: G61–G71.

Culligan PJ and Sinfield JV (2001) Use of NMR relaxation times to differentiate mobile and immobile pore fractions in a wetland soil. *Water Resources Research* 37: 837–842.

Dalgaard E, Auken E, and Larsen J (2012) Adaptive noise cancelling of multichannel magnetic resonance sounding signals. *Geophysical Journal International* 191(1): 88–100.

Dlugosch R, Mueller-Petke M, Guenther T, Costabel S, and Yaramanci U (2011) Assessment of the potential of a new generation of surface NMR instruments. *Near Surface Geophysics* 9: 89–102.

Dlugosch R, Guenther T, Mueller-Petke M, and Yaramanci U (2013) Improved prediction of hydraulic conductivity for coarse-grained, unconsolidated material from nuclear magnetic resonance. *Geophysics* 78: EN55–EN64.

Dunn KJ, Bergman DJ, and Latorraca GA (2002) *Nuclear Magnetic Resonance, Petrophysical and Logging Applications*, first ed. New York: Pergamon.

Fantazzini P, Bortolotti V, Kärger J, and Galvosas P (2010) Editorial for: Magnetic resonance in porous media. *AIP Conference Proceedings*, vol. 133, 1–4.

Flaum M, Chen J, and Hirasaki GJ (2005) NMR diffusion editing for D-T2 maps: Application to recognition of wettability change. *Petrophysics* 46: 113–123.

Freedman R and Heaton N (2004) Fluid characterization using nuclear magnetic resonance logging. *Petrophysics* 45: 241–250.

Freedman R, Heaton N, Flaum M, Hirasaki G, Flaum C, and Hürlimann M (2002) Wettability, saturation, and viscosity using the magnetic resonance fluid characterization method and new diffusion-editing pulse sequences. *SPE Annual Technical Conference and Exhibition*, vol. 13.

Fukushima E and Roeder SB (1981) *Experimental Pulse NMR – A Nuts and Bolts Approach*. Boulder: Westview Press.

Grunewald E and Knight R (2011) The effect of pore size and magnetic susceptibility on the surface NMR relaxation parameter T2*. *Near Surface Geophysics* 9: 169–178.

Grunewald E and Knight R (2012) Nonexponential decay of the surface-NMR signal and implications for water content estimation. *Geophysics* 77: EN1–EN09.

Grunewald E and Walsh D (2013) Multiecho scheme advances surface NMR for aquifer characterization. *Geophysical Research Letters* 40(24): 6346–6350.

Grunewald E, Knight R, and Walsh D (2014) Advancement and validation of surface nuclear magnetic resonance spin-echo measurements of T2. *Geophysics* 79(2): EN15–EN23.

Günther T and Müller-Petke M (2012) Hydraulic properties at the North Sea island of Borkum derived from joint inversion of magnetic resonance and electrical resistivity soundings. *Hydrology and Earth System Sciences* 16: 3279–3291.

Hahn EL (1950) Spin echoes. *Physical Review* 80: 580–594.

Hertrich M, Braun M, and Yaramanci U (2005) Magnetic resonance soundings with separated transmitter and receiver loops. *Near Surface Geophysics* 3: 141–154.

Hertrich M, Braun M, Günther T, Green AG, and Yaramanci U (2007) Surface nuclear magnetic resonance tomography. *IEEE Transactions on Geoscience and Remote Sensing* 45: 3752–3759.

Hertrich M (2008) Imaging of groundwater with nuclear magnetic resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy* 53: 227–248.

Hertrich M, Green AG, Braun M, and Yaramanci U (2009) High-resolution surface-NMR tomography of shallow aquifers based on multi-offset measurements. *Geophysics* 74: G47–G59.

Hübner W, Hunze S, Baumgarten H, Orilski J, and Wonik T (2013) Petrophysical and sedimentary petrographic characterisation of the Bückeburg Formation (German Wealden) in the geothermal well Groß Buchholz Gt-1 (Hanover, Germany). *Zeitschrift der Deutschen Gesellschaft für Geowissenschaften* 163(4): 483–492.

Hürlimann MD, Helmer K, Latour L, and Sotak C (1994) Restricted diffusion in sedimentary rocks. Determination of surface-area-to-volume ratio and surface relaxivity. *Journal of Magnetic Resonance, Series A* 111: 169–178.

Hürlimann MD and Griffin DD (2000) Spin dynamics of Carr–Purcell–Meiboom–Gill-like sequences in grossly inhomogeneous B0 and B1 fields and application to NMR well logging. *Journal of Magnetic Resonance* 143: 120–135.

Hürlimann MD and Venkataramanan L (2002) Quantitative measurement of two-dimensional distribution functions of diffusion and relaxation in grossly inhomogeneous fields. *Journal of Magnetic Resonance* 157: 31–42.

Hürlimann MD, Venkataramanan L, and Flaum C (2002) Diffusion-editing: New NMR measurement of saturation and pore geometry. *SPWLA 43rd Annual Logging Symposium*, vol. 14. .

Jackson JA, Burnett LJ, and Harmon JF (1980) Remote (inside-out) NMR. III. Detection of nuclear magnetic resonance in a remotely produced region of homogeneous magnetic field. *Journal of Magnetic Resonance (1969)* 41: 411–421.

Jiang C, Lin J, Duan Q, Sun S, and Tian B (2011) Statistical stacking and adaptive notch filter to remove high-level electromagnetic noise from MRS measurements. *Near Surface Geophysics* 9: 459–468.

Kenyon WE, Day P, Straley C, and Willemsen J (1988) A three-part study of NMR longitudinal relaxation properties of water-saturated sandstones. *SPE Formation Evaluation* 3: 622–636.

Kenyon WE (1997) Petrophysical principles of applications of NMR logging. *The Log Analyst* 38: 21–43.

Kleinberg RL and Horsfield MA (1990) Transverse relaxation processes in porous sedimentary rock. *Journal of Magnetic Resonance (1969)* 88: 9–19.

Kleinberg RL, Sezginer A, Griffin DD, and Fukuhara M (1992) Novel NMR apparatus for investigating an external sample. *Journal of Magnetic Resonance (1969)* 97: 466–485.

Kleinberg RL, Farooqui SA, and Horsfield MA (1993) T_1/T_2 ratio and frequency dependence of NMR relaxation in porous sedimentary rocks. *Journal of Colloid and Interface Science* 158: 195–198.

Kleinberg RL (1994) Pore size distributions, pore coupling, and transverse relaxation spectra of porous rocks. *Magnetic Resonance Imaging* 12: 271–274, Proceedings of the Second International Meeting on Recent Advances in MR Applications to Porous Media.

Kleinberg RL, Kenyon WE, and Mitra PP (1994) Mechanism of NMR relaxation of fluids in rock. *Journal of Magnetic Resonance, Series A* 108: 206–214.

Kleinberg RL (1996) Utility of NMR T2 distributions, connection with capillary pressure, clay effect, and determination of the surface relaxivity parameter r2. *Magnetic Resonance Imaging* 14: 761–767.

Kozeny J (1927) Über kapillare Leitung des Wassers im Boden. *Sitzungsberichte / Akademie der Wissenschaften in Wien, Mathematisch-Naturwissenschaftliche Klasse Abteilung IIa* 136: 271–306.

Krynicki K, Green CD, and Sawyer DW (1978) Pressure and temperature dependence of self-diffusion in water: Faraday Discuss. *Chemical Society* 66: 199–208.

LaTorraca GA, Dunn KJ, and Brown RJS (1993) Predicting permeability from nuclear magnetic resonance and electrical properties measurements. In: *Society of Core Analysts Conference, Paper 9312*.

Latour LL, Kleinberg RL, and Sezginer A (1992) Nuclear magnetic resonance properties of rocks at elevated temperatures. *Journal of Colloid and Interface Science* 150: 535–548.

Latour LL, Kleinberg RL, Mitra PP, and Sotak CH (1995) Pore-size distributions and tortuosity in heterogeneous porous media. *Journal of Magnetic Resonance, Series A* 112: 83–91.

Legchenko A (2004) Magnetic resonance sounding: Enhanced modeling of a phase shift. *Applied Magnetic Resonance* 25(3): 621–636.

Legchenko A, Baltassat J-M, Bobachev A, Martin C, Robain H, and Vouillamoz J-M (2004) Magnetic resonance sounding applied to aquifer characterization. *Ground Water* 42: 363–373.

Legchenko A, Descloitres M, Vincent C, et al. (2011) Three-dimensional magnetic resonance imaging for groundwater. *New Journal of Physics* 13: 17.

Legchenko A and Shushakov OA (1998) Inversion of surface NMR data. *Geophysics* 63: 75–84.

Legchenko A, Vouillamoz J-M, and Roy J (2010) Application of the magnetic resonance sounding method to the investigation of aquifers in the presence of magnetic materials. *Geophysics* 75: L91–L100.

Legchenko A, Plata JL, and Yaramanci U (2011) Foreword for special topic: Advances in magnetic resonance technology. *Near Surface Geophysics* 9(2): 87.

Lehmann-Horn JA, Hertrich M, Greenhalgh SA, and Green AG (2011) Three-dimensional magnetic field and NMR sensitivity computations incorporating conductivity anomalies and variable-surface topography. *IEEE Transactions on Geoscience and Remote Sensing* 49: 3878–3891.

Levitt MH (2002) *Spin Dynamics – Basics of Nuclear Magnetic Resonance*. Chichester: John Wiley & Sons LTD.

Mansfield P and Morris PG (1982) *NMR Imaging in Biomedicine*. New York: Academic Press.

Meiboom S and Gill D (1958) Modified spin-echo method for measuring nuclear relaxation times. *Review of Scientific Instruments* 29: 688.

Menger S and Sands P (2003) NMR in motion. *Petrophysics* 44: 190–195.

Miller M, Paltiel Z, Gillen M, Granot J, and Bouton J (1990) Spin echo magnetic resonance logging: Porosity and free fluid index determination. *SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers*, vol. 14. .

Mitra PP, Sen PN, Schwartz LM, and Le Doussal P (1992) Diffusion propagator as a probe of the structure of porous media. *Physical Reviews Letters* 68(24): 3555–3558.

Mohnke O and Yaramanci U (2008) Pore size distributions and hydraulic conductivities of rocks derived from magnetic resonance sounding relaxation data using multi-exponential decay time inversion. *Journal of Applied Geophysics* 66: 73–81.

Mueller-Petke M and Yaramanci U (2010) QT inversion—Comprehensive use of the complete surface NMR data set. *Geophysics* 75: WA199–WA209.

Müller-Petke M, Dlugosch R, and Yaramanci U (2011a) Evaluation of surface nuclear magnetic resonance-estimated subsurface water content. *New Journal of Physics* 13: 095002.

Müller-Petke M, Hiller T, Herrmann R, and Yaramanci U (2011b) Reliability and limitations of surface NMR assessed by comparison to borehole NMR. *Near Surface Geophysics* 9: 123–134.

Müller-Petke M and Costabel S (2013) Comparison and optimal parameter setting of reference based harmonic noise cancellation in time and frequency domain for surface-NMR. *Near Surface Geophysics* 12: 199–210. http://dx.doi.org/10.3997/1873-0604.2013041.

Müller-Petke M, Walbrecker JO, and Knight R (2013) The inversion of surface-NMR T_1 data for improved aquifer characterization. *Geophysics* 78: EN83–EN94.

Pape H, Tillich JE, and Holz M (2006) Pore geometry of sandstone derived from pulsed field gradient NMR. *Journal of Applied Geophysics* 58: 232–252.

Parsekian AD, Grosse G, Walbrecker JO, et al. (2013) Detecting unfrozen sediments below thermokarst lakes with surface nuclear magnetic resonance. *Geophysical Research Letters* 40: 535–540. http://dx.doi.org/10.1002/grl.50137.

Plata JL and Rubio FM (2008) The use of MRS in the determination of hydraulic transmissivity: The case of alluvial aquifers. *Journal of Applied Geophysics* 66: 128–139 (resonance sounding – a reality in Applied Hydrogeophysics).

Prammer MG, Drack E, Goodman G, et al. (2000) The magnetic-resonance while-drilling tool: Theory and operation. In: *75th Annual Technical Conference and Exhibition: Society of Petroleum Engineers,* Paper SPE-62981.

Purcell EM, Torrey HC, and Pound RV (1946) Resonance absorption by nuclear magnetic moments in a solid. *Physical Review* 69: 37–38.

Radic T (2006) Improving the signal-to-noise ratio of surface NMR data due to the remote reference technique. In: *Presented at the 12th European Meeting of Environmental and Engineering Geophysics* .

Sapoval B, Russ S, Petit D, and Korb J (1996) Fractal geometry impact on nuclear relaxation in irregular pores. *Magnetic Resonance Imaging* 14: 863–867.

Schirov MD, Legchenko A, and Creer G (1991) A new direct non-invasive groundwater detection technology for Australia. *Exploration Geophysics* 22: 333–338.

Schoenfelder W, Glser HR, Mitreiter I, and Stallmach F (2008) Two-dimensional NMR relaxometry study of pore space characteristics of carbonate rocks from a Permian aquifer. *Journal of Applied Geophysics* 65: 21–29.

Seevers D (1966) A nuclear magnetic method for determining the permeability of sandstones. In: *SPWLA 7th Annual Logging Symposium* 1–4.

Semenov A (1987) NMR hydroscope for water prospecting. In: *Proceedings of the Seminar on Geotomography* 66–67.

Semenov A, Schirov M, Legchenko A, Burshtein A, and Pusep A (1989) *Device for Measuring the Parameter of Underground Mineral Deposits*, Patent 2198540B (Great Britain).

Simpson JH and Carr HY (1958) Diffusion and nuclear spin relaxation in water. *Physical Review* 111: 1201–1202.

Simunek J, van Genuchten MT, and Sejna M (1999) Using the HYDRUS-1D and HYDRUS-2D codes for estimating unsaturated soil hydraulic and solute transport parameters. In: *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*, 1523–1536.

Song Y-Q (2003) Using internal magnetic fields to obtain pore size distributions of porous media. *Concepts in Magnetic Resonance Part A* 18A: 97–110.

Song Y-Q, Venkataramanan L, Hürlimann MD, Flaum M, Frulla P, and Straley C (2002) T1–T2 correlation spectra obtained using a fast two-dimensional Laplace inversion. *Journal of Magnetic Resonance* 154: 261–268.

Song Y-Q (2012) Focus on the physics of magnetic resonance on porous media. *New Journal of Physics* 14: 055017.

Stejskal EO and Tanner JE (1965) Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics* 42: 288.

Straley C, Rossini D, Vinegar H, Tutunjian P, and Morriss C (1997) Core analysis by low-field NMR. *The Log Analyst* 38: 84–94.

Timur A (1968) An investigation of permeability, porosity, and residual water saturation relationships for sandstone reservoirs. *The Log Analyst* IX(4): 8–17.

Timur A (1969) Pulsed magnetic resonance studies of porosity, movable fluid and permeability of sandstones. *Journal of Petroleum Technology* 21: 775–786.

Valla P and Yaramanci U (2002) Foreword from the guest editors for: Surface nuclear magnetic resonance: What is possible? *Journal of Applied Geophysics* 50(1–2): 1–2.

Varian R (1962) *Ground Liquid Prospecting Method and Apparatus*, US Patent 3019383.

Vogt C, Galvosas P, Klitzsch N, and Stallmach F (2002) Self-diffusion studies of pore fluids in unconsolidated sediments by PFG NMR. *Journal of Applied Geophysics* 50: 455–467.

Walbrecker JO, Hertrich M, and Green AG (2009) Accounting for relaxation processes during the pulse in surface NMR data. *Geophysics* 74: G27–G34.

Walbrecker JO, Hertrich M, and Green AG (2011a) Off-resonance effects in surface nuclear magnetic resonance. *Geophysics* 76: G1–G12.

Walbrecker JO, Hertrich M, Lehmann-Horn JA, and Green AG (2011b) Estimating the longitudinal relaxation time T_1 in surface NMR. *Geophysics* 76: F111–F122.

Walsh DO (2008) Multi-channel surface NMR instrumentation and software for 1D/2D groundwater investigations. *Journal of Applied Geophysics* 66: 140–150.

Walsh DO, Grunewald E, Turner P, Hinnell A, and Ferre P (2011) Practical limitations and applications of short dead time surface NMR. *Near Surface Geophysics* 9: 103–111.

Walsh D, Turner P, Grunewald E, et al. (2013) A small-diameter NMR logging tool for groundwater investigations. *Groundwater* 51: 914–926. http://dx.doi.org/10.1111/gwat.12024.

Weichman PB, Lavely EM, and Ritzwoller MH (2000) Theory of surface nuclear magnetic resonance with applications to geophysical imaging problems. *Physical Review E* 62: 1290–1312.

Whittall KP, Bronskill MJ, and Henkelman R (1991) Investigation of analysis techniques for complicated NMR relaxation data. *Journal of Magnetic Resonance (1969)* 95: 221–234.

Winkler M, Freeman JJ, and Appel M (2005) The limits of fluid property correlations used in NMR well logging: An experimental study of reservoir fluids at reservoir conditions. *Petrophysics* 46: 104–112.

Yaramanci U and Legchenko A (2005) Guest editorial for special topic: Magnetic resonance sounding. *Near Surface Geophysics* 3(3): 119–120.

Yaramanci U, Legchenko A, and Roy J (2008) Magnetic resonance sounding. *Journal of Applied Geophysics* 66(3–4): 71–72 (special issue).

Yaramanci U, Knight R, and Legchenko A (2014) Foreword for special topic: Magnetic Resonance in the Subsurface. *Near Surface Geophysics* 12(2): 183.

Zhang GQ, Huang C, and Hirasaki GJ (2000) Interpretation of wettability in sandstones with NMR analysis. *Petrophysics* 41: 223–233.

This page intentionally left blank

11.14 Tools and Techniques: Radiometric Methods

PG Killeen, Geophysical Consultant (Retired Research Scientist, Geological Survey of Canada)
CJ Mwenifumbo, Consulting Geophysicist (Retired Research Scientist, Geological Survey of Canada)
KL Ford, Geologist (Retired Physical Scientist, Geological Survey of Canada)

© 2015 Elsevier B.V. All rights reserved.

11.14.1 Introduction 449
11.14.1.1 Brief Historical Background 449
11.14.1.2 Earth Science and Gamma Rays 449
11.14.2 Fundamental Principles and Theory 450
11.14.2.1 Radioactivity 450
11.14.2.2 Natural Radioactive Elements in Soil and Rocks 451
11.14.2.3 The U and Th Radioactive Decay Series and Potassium 452
11.14.2.3.1 Uranium 452
11.14.2.3.2 Thorium 452
11.14.2.3.3 Potassium 452
11.14.2.4 Radioactive Equilibrium 452
11.14.2.4.1 Radon measurement as a radiometric technique 454
11.14.2.4.2 Equivalent uranium and equivalent thorium 454
11.14.2.5 The Theoretical Gamma Ray Energy Spectrum 455
11.14.2.6 Interaction of Gamma Rays with Geologic Materials 455
11.14.2.7 Detecting and Counting Gamma Rays 456
11.14.2.7.1 Geiger–Müller counters 456
11.14.2.7.2 Scintillometers and scintillation detectors 456
11.14.2.7.3 Gamma ray spectrometers 457
11.14.2.8 The Measured Gamma Ray Energy Spectrum 458
11.14.2.8.1 Stripping factors 459
11.14.2.9 Source–Detector Geometry 460
11.14.2.9.1 Borehole measurement geometry 460
11.14.2.9.2 Surface measurement geometry 460
11.14.2.9.3 Airborne measurement geometry 461
11.14.2.10 Effective Sample Volume 461
11.14.2.10.1 Borehole measurement sample volume 461
11.14.2.10.2 Surface measurement sample volume 462
11.14.2.10.3 Airborne measurement sample volume 462
11.14.3 Instrumentation 463
11.14.3.1 Introduction 463
11.14.3.2 Counting Statistics 463
11.14.3.3 Instrument Parameters in GRS 463
11.14.3.3.1 Detectors used in Earth sciences 463
11.14.3.3.2 Stabilization 464
11.14.3.3.3 Pulse pileup 464
11.14.3.3.4 Spectrum linearity 465
11.14.3.3.5 Dead time 465
11.14.3.4 Some Features of New Systems 465
11.14.3.5 Calibration for Quantitative Radioelement Measurements 466
11.14.3.6 Calibration Facilities 466
11.14.3.6.1 Introduction 466
11.14.3.6.2 Calibration pads and model boreholes 467
11.14.3.6.3 Construction of calibration facilities 467
11.14.4 Data Acquisition 468
11.14.4.1 Introduction 468
11.14.4.2 Data Acquisition with Surface Systems 468
11.14.4.2.1 Portable gamma ray spectrometers 468
11.14.4.2.2 Car-borne gamma ray spectrometers 469
11.14.4.2.3 Background radiation in surface measurements 469
11.14.4.2.4 Calibration of portable gamma ray spectrometers 469
11.14.4.2.5 Determination of sensitivities for car-borne and airborne systems 470
11.14.4.3 Data Acquisition with Airborne Systems 470

11.14.4.3.1 Introduction 470
11.14.4.3.2 Airborne gamma ray spectrometric surveying 470
11.14.4.3.3 The airborne gamma ray spectrometer requirements 470
11.14.4.3.4 Calibration of airborne gamma ray spectrometers 471
11.14.4.3.5 Determination of attenuation coefficients 471
11.14.4.3.6 Determination of window sensitivities 472
11.14.4.3.7 Airborne gamma ray spectrometer field verification tests 472
11.14.4.4 Data Acquisition with Borehole Logging Systems 472
11.14.4.4.1 Introduction 472
11.14.4.4.2 TC gamma ray and gamma ray spectral logging 473
11.14.4.4.3 Background radiation in borehole measurements 473
11.14.4.4.4 Determination of borehole correction factors 473
11.14.4.4.5 Calibration of natural gamma ray logging systems 474
11.14.5 Data Processing 475
11.14.5.1 Introduction 475
11.14.5.2 Data Processing for Surface Systems 475
11.14.5.2.1 Portable gamma ray spectrometers 475
11.14.5.2.2 Car-borne gamma ray spectrometers 476
11.14.5.3 Data Processing for Airborne Gamma Ray Spectrometer Systems 476
11.14.5.3.1 Noise reduction 476
11.14.5.3.2 AGRS windowing procedure and energy calibration 477
11.14.5.3.3 Dead time correction 477
11.14.5.3.4 Filtering to prepare for background corrections 477
11.14.5.3.5 Cosmic and aircraft background 477
11.14.5.3.6 Radon background 477
11.14.5.3.7 Calculation of effective height above ground level 479
11.14.5.3.8 Stripping 479
11.14.5.3.9 Attenuation correction 479
11.14.5.3.10 Computation of apparent radioelement concentrations on the ground 479
11.14.5.3.11 Calculation of radioelement ratios 479
11.14.5.4 Data Processing for Borehole Systems 479
11.14.5.4.1 TC gamma ray logs 479
11.14.5.4.2 Gamma ray spectral logs 481
11.14.6 Data Presentation and Visualization 482
11.14.6.1 Introduction 482
11.14.6.2 Variables and Standard Survey Products for AGRS Data 482
11.14.6.2.1 Digital data 482
11.14.6.2.2 Gridding and color interval maps 482
11.14.6.2.3 Line data presentation 483
11.14.6.2.4 Cross plots for surface and airborne gamma ray spectrometry data 484
11.14.6.3 Visual Enhancement of AGRS Data 486
11.14.6.3.1 Single band 486
11.14.6.3.2 Multiband 486
11.14.6.3.3 Enhancement 486
11.14.6.3.4 Integrated images 486
11.14.6.3.5 Dimensional visualizations 490
11.14.6.4 Data Presentation and Visualization of BGRS Data 490
11.14.7 Data Analysis 490
11.14.7.1 Introduction 490
11.14.7.2 Classification 491
11.14.7.2.1 Unsupervised classification 492
11.14.7.2.2 Supervised classification 492
11.14.7.3 Analysis of BGRS Data 496
11.14.7.3.1 TC and spectral log data analysis 496
11.14.7.3.2 TC and spectral log data interpretation 496
11.14.8 Interpretation and Applications 497
11.14.8.1 Introduction 497
11.14.8.2 General Comments on Interpretation 497
11.14.8.3 Bedrock Geologic Mapping 499
11.14.8.4 Surficial Geologic Mapping 502
11.14.8.5 Mineral Exploration 503

11.14.8.5.1 Direct detection: uranium 504
11.14.8.5.2 Direct detection: U and Th as trace elements 505
11.14.8.5.3 Indirect detection: alteration mapping 506
11.14.8.6 Interpretation and Applications of BGRS 509
11.14.8.6.1 Geologic mapping: definition of geologic contacts 509
11.14.8.6.2 Stratigraphic correlation 511
11.14.8.6.3 Mineral exploration 512
11.14.8.6.4 Kimberlite characteristics and classification: diamond exploration 514
11.14.8.7 Environmental Radiation Monitoring 514
11.14.8.7.1 Radon potential mapping 515
11.14.8.7.2 Mapping fallout from nuclear incidents 515
11.14.8.7.3 Monitoring nuclear power facilities 517
11.14.8.7.4 Emergency radiation response and preparedness 517
11.14.9 Conclusions and Future Directions 518
Acknowledgments 519
References 520

Abbreviations

ADC	Analog to digital converter
AGRS	Airborne gamma ray spectrometry/spectrometer
AGSO	Australian Geological Survey Organization
API	American Petroleum Institute
ASL	Above sea level
BGO	Bismuth germanate
BGRS	Borehole gamma ray spectrometry
cps	Counts per second
CsI(Na)	Cesium iodide (sodium-activated)
DEM	Digital elevation model
eU, eTh	Equivalent U, equivalent Th
FOV	Field of view
FWHM	Full width (at) half maximum
GGRS	Ground gamma ray spectrometry/spectrometer
GIS	Geographic information system
GPS	Global Positioning System
GRS	Gamma ray spectrometry
GSC	Geological Survey of Canada
IAEA	International Atomic Energy Agency
IOCG	Iron oxide copper-gold
LLD	Low-level discriminator
MNF	Maximum noise fraction
NADR	Natural air absorbed dose rate
NaI(Tl)	Sodium iodide (thallium-activated)
NASVD	Noise-adjusted singular value decomposition
NURE	National Uranium Resource Evaluation (US program)
PCA	Principal component analysis
PHA	Pulse height analyzer
PMT	Photomultiplier tube
RA log	Radiometric assay log
REE	Rare earth element
RPM	Remote predictive mapping
SCA	Single-channel analyzer
SRTM	Shuttle Radar Topographic Mission
STP	Standard temperature and pressure
TADR	Total air absorbed dose rate
TC	Total count
UN	United Nations
USDOE	United States Department of Energy
VMS	Volcanogenic massive sulfide
XRF	x-Ray fluorescence

Symbols

A	Atomic mass
Z	Atomic number
Z_{eq}	Equivalent Z

11.14.1 Introduction

11.14.1.1 Brief Historical Background

Nuclear radiation was first observed by Henri Becquerel in 1896 when he noticed the darkening of photographic plates in a drawer, which also contained a specimen of uranium sulfate. This led to the discovery of the phenomenon, which Marie Curie, in 1898, gave the name 'radioactivity.' Her work was the basis for the 'radiometric techniques.' The brief review of the physics of the atom given in the succeeding text defines the terminology and origins of the different types of nuclear radiation, which are fundamental to the radiometric techniques used in geophysical surveys.

11.14.1.2 Earth Science and Gamma Rays

This chapter with Section 11.14.2 covers the natural radiation from rocks and minerals and the measurement of that

radiation. This section provides the foundation for later detailed Sections 11.14.3–11.14.5, 11.14.7, and 11.14.8. In particular, it describes how useful geologic information can be obtained by measuring natural gamma radiation. The concept of using gamma ray surveys to detect potassium (K), uranium (U), and thorium (Th) variations associated with mineral deposits is introduced.

The fundamental information needed to effectively apply gamma ray spectrometric techniques is provided, including the factors that affect measurements such as the interaction of gamma rays with matter. The meaning of 'radioactive equilibrium' in the uranium decay series in nature and its importance as a basic assumption of gamma ray spectrometry (GRS) are described, along with a discussion of the geometry of the measurements, the effective sample volume, and background radiation.

The detection of radiation and the associated instrumentation for data acquisition are described starting with the basics of gamma ray measurements with Geiger counters, scintillometers, and gamma ray spectrometers. It includes the properties of detectors, their energy resolution, and the important characteristics of gamma ray spectrometers such as dead time, which affect results, as well as the basics of gamma ray spectrometer calibration.

With the previously mentioned fundamentals as background information, the detailed sections follow, first with instrumentation. This includes a brief note about counting statistics, followed by description of specifications and characteristics for today's state-of-the-art instrumentation for gamma ray and multichannel gamma ray spectrometric systems in airborne, surface, and borehole surveys. A description of existing calibration facilities for quantitative measurements from the air, on the ground, and in boreholes is also given. This report follows the recommendations of the International Atomic Energy Agency (IAEA). The IAEA is a United Nations organization with the stated objective "to accelerate and enlarge the contribution of atomic energy to peace, health and prosperity throughout the world" (IAEA, 1991).

Data acquisition is then described along with surveying methodology. The acquisition of calibration data and calibration procedures is also covered. Radon measurement, another radiometric survey technique, is covered briefly. This is followed by descriptions of the methods for processing, presentation, analysis, and interpretation of data.

State-of-the-art information regarding the application and interpretation of radiometric geophysical techniques in Earth sciences with emphasis on gamma ray spectrometric methods follows. It includes geologic problems (e.g., mapping and mineral exploration, both direct and indirect) and environmental problems (e.g., contamination, waste sites (landfill), and radon). Examples of applications are illustrated with various types of survey results, presentation formats, and interpretations. The reviews of applications and interpretation include airborne and ground (surface) gamma ray surveys and borehole gamma ray logging.

11.14.2 Fundamental Principles and Theory

11.14.2.1 Radioactivity

Every chemical element in the periodic table is composed of atoms with a nucleus surrounded by orbiting electrons. The nucleus is composed of protons (positively charged) and neutrons (no charge), both with essentially the same mass. In comparison, electrons (negatively charged) have almost negligible mass.

Each element is defined by its atomic number Z, which is the number of protons in the nucleus. The atomic mass A is the sum of the masses of the protons and neutrons. An element E is denoted by its chemical symbol, atomic mass A, and atomic number Z as follows: ${}^{A}E_{Z}$, for example, ${}^{238}U_{92}$. In physics literature, the more general term 'nuclide' is used to refer to every unique combination of protons and neutrons and permits the complete description of every known atomic nucleus, either stable or radioactive. When the number of neutrons is larger or smaller than the number of protons in an atom of an element, it is still the same element with the same chemical properties but is called an isotope of that element with a different atomic mass. An element may have one or more stable or long-lived isotopes. Some isotopes (or nuclides) have excess energy in the nucleus, which makes them unstable. Then, they are called radioisotopes (or radionuclides), often referred to as the parent, which disintegrate into a more stable (usually) nuclide, often called the daughter product. This process is called radioactive decay, in which the excess energy is reduced by the emission of radiation in the form of alpha particles, beta particles, or gamma rays. The number N of atoms of a decaying radioisotope, at any time t, is given by

$$N_t = N_0 e^{-\lambda t} \quad [1]$$

where N_0 is the number of atoms at time $t=0$ and λ is the decay constant for that radioisotope with units of s^{-1}. The detection of radioactivity and the measurement and counting of alpha particles, beta particles, and especially gamma rays form the basis of the radiometric techniques used in Earth sciences.

An alpha particle is essentially a helium nucleus stripped of electrons, which resembles a doubly ionized helium atom with an atomic mass of 4 and a double-positive charge. Beta particles are essentially the same as electrons, the only difference being their place of origin. Alpha and beta particles are emitted from the nucleus of a radioactive element (or radioelement). Electrons originate from an atomic orbit around the nucleus. Gamma rays are photons emitted from the nucleus after a radioactive decay by either alpha or beta emission (or by electron capture), when the nucleus of the daughter element is in an excited state. The gamma radiation represents the energy emitted as the excited daughter nucleus goes to the ground state. Because gamma rays have no mass or charge, they can penetrate a significant thickness of materials compared to alpha particles, which have both mass and electrical charge and electrons, which have a charge. Gamma rays originating from natural radioactive sources in about the top 50 cm of rocks and soils have energies ranging from 0 to 3.0 MeV (million electron volts) and can be detected and measured at distances of over 700 m in air making gamma ray detection in the field a practical geophysical survey technique. Alpha particle and beta particle detection and measurement are mainly limited to laboratory measurements or field methods with detectors virtually in direct contact with the source such as in some radon gas measurement techniques.

Both alpha particles and gamma rays are emitted with discrete (specific) energies, which are characteristic of the

decaying radioelement. Measurement of the energies of these radiations can be used to identify and determine the amount of the naturally occurring radioactive elements in geologic materials. This forms the basis of alpha particle spectrometry and GRS. Alpha particle spectrometry is basically a laboratory analysis technique because the samples must be counted in a vacuum chamber so that air does not attenuate the alpha particles. However, an alpha particle spectrometer has been used on the moon as a 'field' instrument in the 'Lunar Prospector' operated by the National Aeronautic and Space Administration. Beta particles are not emitted with discrete energies but have a continuous distribution of energies with a maximum energy, which is unique to the radioelement. General descriptions of radioactivity are given by Lapp and Andrews (1963) and Kaplan (1964), while more detailed information is given by Evans (1955), Siegbahn (1968), and Tsoulfanidis and Landsberger (2010).

11.14.2.2 Natural Radioactive Elements in Soil and Rocks

Some important radioactive elements are found in nature. Those of ^{238}U and its radioactive decay products and ^{232}Th and its decay products produce detectable radioactivity. Another isotope of uranium in nature, ^{235}U, which makes up 0.72% of natural uranium, is relatively unimportant in radiometric techniques except in a few situations. Average crustal abundance of uranium and thorium is approximately 2.5 and 10 ppm, respectively. A third important radioelement is the potassium isotope, ^{40}K. All of these are found in varying concentrations in most rock-forming minerals. Although ^{40}K represents only 0.0118% of natural potassium, it contributes significant radiation because of the relatively high abundance of potassium in rocks (approximately 2.5% of the Earth's crust; IAEA, 2003). Potassium occurs mainly in the mineral feldspar, which is an abundant and widespread mineral in the Earth's crust. It also occurs in other common rock-forming minerals, such as in mica, biotite, and muscovite. Uranium and thorium are generally present in low concentrations measured in parts per million (ppm) in a wide range of minerals. The range of radioactive element concentrations in most common rock types is shown in **Table 1** (Killeen, 1979).

Originally, knowledge of the distribution of radioelements was considered useful primarily for uranium exploration (IAEA, 1979) and for computing radiogenic heat production (Killeen and Richardson, 1980; Sass et al., 1968). However, the three radioactive elements can provide an indication of economic concentrations of many other metals besides uranium. They can be indirect indicators of other types of mineralization if they occur as associated trace elements or if their enrichment or depletion is due to geochemical alteration related to economic mineralization. For example, potassic alteration is associated with certain mineral deposit types, including porphyry copper (Cu–Au–Mo) deposits, volcanogenic massive sulfide (VMS) deposits, and iron oxide copper-gold (IOCG) deposits (Shives et al., 1997). Maps that show the radioactive element distribution provide information about fundamental mineralogical and geochemical properties of bedrock and surficial deposits. These are useful for geologic mapping, mineral exploration, and, in addition, environmental radiation monitoring (Ford et al., 2008a,b). Today, radiometric techniques have evolved to the point where the distribution of the radioelements can be mapped with airborne survey methods.

Table 1 Radioelement concentrations in different classes of rocks

Rock type	*Potassium (%)*		*Uranium (ppm)*		*Thorium (ppm)*	
	Mean	*Range*	*Mean*	*Range*	*Mean*	*Range*
Felsic extrusives	3.1	1.0–6.2	4.1	0.8–16.4	11.9	1.1–41.0
Felsic intrusives	3.4	0.1–7.6	4.5	0.1–30.0	25.7	0.1–253.1
Intermediate extrusives	1.1	1.1–2.5	1.1	0.2–2.6	2.4	0.4–6.4
Intermediate intrusives	2.1	0.1–6.2	3.2	0.1–23.4	12.2	0.4–106.0
Mafic extrusives	0.7	0.06–2.4	0.8	0.03–3.3	2.2	0.05–8.8
Mafic intrusives	0.8	0.01–2.6	0.8	0.01–5.7	2.3	0.03–15.0
Ultramafic	0.3	0–0.8	0.3	0–1.6	1.4	0–7.5
Alkali feldspathoidal intermediate extrusives	6.5	2.0–9.0	29.7	1.9–62.0	133.9	9.5–265.0
Alkali feldspathoidal intermediate intrusives	4.2	1.0–9.9	55.8	0.3–720.0	132.6	0.4–880.0
Alkali feldspathoidal mafic extrusives	1.9	0.2–6.9	2.4	0.5–12.0	8.2	2.1–60.0
Alkali feldspathoidal mafic intrusives	1.8	0.3–4.8	2.3	0.4–5.4	8.4	2.8–19.6
Chemical sedimentary rocks	0.6	0.02–8.4	3.6	0.03–26.7	14.9	0.03–132.0
Carbonates	0.3	0.01–3.5	2	0.03–18.0	1.3	0.03–10.8
Detrital sedimentary rocks	1.5	0.01–9.7	4.8	0.01–80.0	12.4	0.2–362.0
Metamorphosed igneous rocks	2.5	0.1–6.1	4	0.1–148.5	14.8	0.1–104.2
Metamorphosed sedimentary rocks	2.1	0.01–5.3	3	0.1–53.4	12	0.1–91.4

Source: Killeen PG (1979) Gamma ray spectrometric methods in uranium exploration: Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 163–230. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

11.14.2.3 The U and Th Radioactive Decay Series and Potassium

11.14.2.3.1 Uranium

Uranium decays in a series starting with the long-lived parent ^{238}U with half-life 4.5 billion years (Table 2). After 14 successive decays of daughter products with half-lives ranging from seconds, to days, to years, to thousands of years, the series ends with a stable isotope of lead (^{206}Pb). In many of these steps, gamma rays are emitted by daughters with energies specific to the particular decaying radioisotope. ^{238}U does not emit gamma rays, which could be detected and used to determine the uranium content. However, some daughters in the decay series do emit gamma rays that are useful for that purpose, in particular ^{214}Pb and ^{214}Bi, the eighth and ninth daughter products in the series. Because there is usually a 'radioactive equilibrium' relationship (discussed later) between the daughter products and the parent, measurement of gamma radiation from the daughters can be used indirectly to determine the concentration of the parent uranium. However, the assumption of radioactive equilibrium (also called secular equilibrium) should be checked in different geologic environments to ensure that the use of this indirect method is valid.

11.14.2.3.2 Thorium

Thorium also decays in a series starting with the long-lived parent ^{232}Th with half-life 13.9 billion years and, after ten successive decays (Table 3), ends with a stable isotope of lead (^{208}Pb). Similar to the uranium series, many of these steps result in the emission of gamma rays with energies specific to the particular decaying radioisotope. Like uranium, the parent ^{232}Th does not emit gamma rays itself but the thorium content of geologic materials can be determined indirectly as in the case of uranium by detecting gamma rays from thallium (^{208}Tl), the ninth daughter product in the series. Unlike the uranium decay series, which has many daughter products with relatively long half-lives ranging up to thousands of years, the longest half-life of any daughter in the thorium series is that of radium (^{228}Ra), which is only 6.7 years. This has ramifications for the assumption of radioactive equilibrium discussed in the succeeding text.

11.14.2.3.3 Potassium

The natural radioactive isotope of potassium (^{40}K) decays with half-life 1.3 billion years into the stable isotope of argon (^{40}Ar) directly (without a decay series) with the emission of gamma rays with energy of 1.46 MeV, which can be detected and used to determine the potassium content directly.

Table 2 Decay series of ^{238}U showing radiation type and half-lives

Isotope	Radiation	Half-life
^{238}U	α	4.507×10^{9} years
↓		
^{234}Th	β	24.1 days
↓		
^{234}Pa	β	1.18 months
↓		
^{234}U	α	2.48×10^{5} years
↓		
^{230}Th	α	7.52×10^{4} years
↓		
^{226}Ra	α	1600 years
↓		
^{222}Rn	α	3.825 days
↓		
^{218}Po	α	3.05 months
↓		
^{214}Pb	β	26.8 months
↓		
^{214}Bi	β	19.7 months
↓		
^{214}Po	α	1.58×10^{-4} s
↓		
^{210}Pb	β	22.3 years
↓		
^{210}Bi	β	5.02 days
↓		
^{210}Po	α	138.4 days
↓		
^{206}Pb	Stable	

Isotopes constituting $<0.2\%$ of the decay products are omitted.

Source: Grasty RL (1979) Gamma ray spectrometric methods in uranium exploration – Theory and operation procedures. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 147–162. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

11.14.2.4 Radioactive Equilibrium

Radioactive equilibrium or disequilibrium is an important consideration in gamma ray spectrometric measurements of radioelements in nature. GRS can be used to determine the

Table 3 Decay series of ^{232}Th showing radiation type and half-lives

Isotope		Radiation	Half-life
^{232}Th		α	1.39×10^{10} years
↓			
^{228}Ra		β^-	6.7 years
↓			
^{228}Ac		β^-	6.13 h
↓			
^{228}Th		α	1.91 years
↓			
^{224}Ra		α	3.64 days
↓			
^{220}Rn		α	55.3 s
↓			
^{216}Po		α	0.158 s
↓			
^{212}Pb		β^-	10.64 h
↓			
^{212}Bi	36%	β^- (64%) α (36%)	60.5 months
^{212}Po		α	3.04×10^{-7} s
64% ^{208}Tl		β^-	3.1 months
↓			
^{208}Pb		Stable	

Source: Grasty RL (1979) Gamma ray spectrometric methods in uranium exploration – Theory and operation procedures. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 147–162. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

concentrations of uranium, thorium, and potassium in rocks and soils because gamma rays of specific energies are associated with each radioelement. By looking at peaks in the energy spectrum of gamma rays being emitted by the source, the radioelement content of the source can be determined. The method involves the counting of gamma ray photons with specified energies, usually those emitted by daughter products, bismuth (^{214}Bi) in the ^{238}U decay series (Table 2) and thallium (^{208}Tl) in the ^{232}Th decay series shown in Table 3. The gamma ray count rate can then be related to the amount of parent, by assuming there is a direct relation between the amount of daughter and that of parent. This assumption is valid when the radioactive decay series is in a state of secular equilibrium.

A radioactive decay series such as that of ^{238}U is said to be in a state of secular equilibrium when the number of atoms of each daughter being produced in the series is equal to the number of atoms of that daughter being lost by radioactive decay.

The rate of loss by decay is proportional to the amount (N) of radioelement present, for example,

$$\frac{dN_1}{dt} = -\lambda_1 N_1 \qquad [2]$$

where N_1 is the amount of element 1 and λ_1 is the decay constant for the element 1.

In a radioactive decay series, N_1 is decaying into N_2 at the rate in the preceding text while at the same time N_2 is decaying with the decay constant λ_2 into N_3 and so on. The half-life ($T_{1/2}$) is the time (in seconds) for half of the radioelement to decay (IAEA, 2003):

$$T_{1/2} = 0.693\lambda^{-1} \qquad [3]$$

The activity of the radioelement is the product λN, in becquerels (Bq). If the parent has a relatively long half-life, after a long period of time, the amount of any given daughter becomes constant. The rate of production from its parent is equal to its activity or rate of decay. The series is then in a state of secular equilibrium.

For a radioactive decay series, secular equilibrium implies that the activities of the daughters are all the same:

$$\lambda_1 N_1 = \lambda_2 N_2 = \lambda_3 N_3 = \ldots = \lambda_n X_n \qquad [4]$$

After a sufficiently long period of time depending on the half-lives of the daughters, when this condition is met, it is possible to determine the amount of the parent of the decay series by measuring the radioactivity from any daughter element.

The assumption of secular equilibrium required for analysis by gamma ray spectrometric techniques is valid for most geologic materials except in certain circumstances. If one or more of the daughter products are being lost by any process other than radioactive decay, or if the parent was not deposited too long ago, eqn [4] will not be satisfied. Since each daughter product is an element with its own characteristic physical and chemical properties, it may behave differently in different geologic environments. The solubilities of radium, uranium, and thorium differ, and preferential leaching of elements may occur. Also, for example, in the ^{238}U decay system, there is a gas that may escape, radon (^{222}Rn), with a 3.85-day half-life.

One of the most complete references on radioactive disequilibrium is by Rosholt (1959) who subdivided the uranium decay series into five separate groups of isotopes with respect to their state of radioactive equilibrium as shown in Figure 1. The elements within each group tend to remain in equilibrium with the longer-lived parent of the group, although the parent of the group may not itself be in equilibrium with the parent of the decay series, ^{238}U.

The length of time for secular equilibrium to become established is dependent on the half-life of the longest-lived daughter in the decay chain below the parent. After a time of one half-life, the daughter will have accumulated one-half (50%) of the amount to be in equilibrium with the parent. Another half-life adds 25% of the equilibrium amount for a total of 75%. A third half-life adds 12.5% and so on until 99.2% of the equilibrium amount is reached after seven half-lives. For most practical purposes, it can be assumed that after seven half-lives (or even less in some cases), secular equilibrium is established.

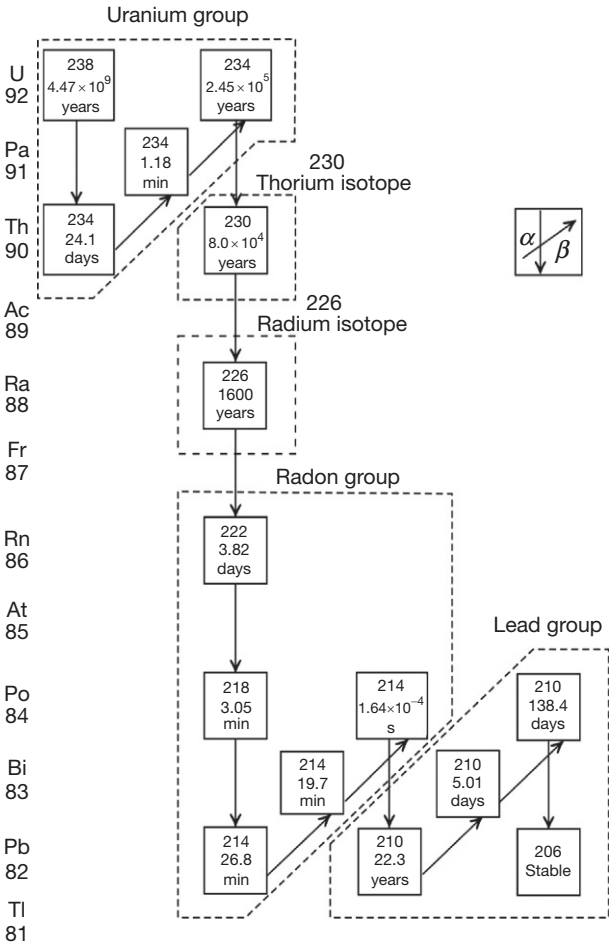


Figure 1 Five equilibrium groupings of isotopes in the decay series of ^{238}U. Within each group, the shorter-lived isotopes tend to remain in equilibrium with the longer-lived parent of the group. This grouping helps to estimate the time to reestablish equilibrium when it is disturbed in nature. Reproduced from Killeen PG (1979) Gamma ray spectrometric methods in uranium exploration: Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 163–230. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

The thorium decay series can generally be assumed to be in radioactive equilibrium. The longest-lived daughter in the thorium series is ^{228}Ra with a half-life of 6.7 years. Seven half-lives, totaling <50 years, is a geologic 'instant,' and any redistribution of thorium in nature would be followed by a relatively rapid reestablishment of secular equilibrium. Thus, in practically all geologic samples, the amount of parent ^{232}Th can be computed by measuring the 2.62 MeV gamma ray activity of the daughter ^{208}Tl.

Some examples of disequilibrium in the uranium decay series are given in the succeeding text.

Example 1. If uranium in solution moves into a chemically reducing environment such as a swamp and accumulates in substantial amounts over a short time, it will be relatively undetectable by gamma ray counting for considerable time (^{238}U does not emit gamma rays). The daughter products must be given time to build up into their equilibrium proportions. If it is assumed (and it reasonably can be) that U isotopes (e.g., the ^{234}U daughter) travel in approximately equilibrium proportions, then the length of time will be controlled by the next daughter in the decay series, ^{230}Th with a half-life of ~80 000 years. The daughters below ^{230}Th have shorter half-lives and will remain in equilibrium with the ^{230}Th as it builds up from decay of the ^{234}U. Fifty percent of the equilibrium amount of ^{230}Th will be attained in one half-life (80 000 years), 75% after two half-lives, and 99.2% after seven half-lives (~560 000 years). For most practical purposes, GRS can be used to determine the amount of parent U by counting 1.76 MeV gamma rays from the daughter ^{214}Bi after about 300 000 years (four half-lives; ~94% of equilibrium). ^{214}Bi is usually chosen because some of its gamma rays are of high energy (and hence are more penetrating), and in addition, there are fewer gamma rays with similar energies emitted by other daughter isotopes, which might interfere with the measured spectrum.

Example 2. Calibration facilities for quantitative gamma ray spectrometric measurements in the field (airborne, surface, and borehole) depend on lab analyses of samples from the calibration models as described in **Section 11.14.3.6**. Disequilibrium can occur in rock samples that are crushed for analysis by a laboratory gamma ray spectrometer. The crushing operation releases some radon and the sample must be placed in a sealed container and allowed to return to equilibrium. In this case, it is the radon group (^{222}Rn, ^{216}Po, ^{214}Pb, ^{214}Bi, and ^{214}Po) defined by Rosholt (1959), which is out of equilibrium with the rest of the decay series. The ^{210}Pb (half-life = 22 years) below the radon group will not be affected by the missing radon if only a short time elapses before the sample is sealed. Again, in about seven half-lives of the longest-lived member of the radon group (3.85 days; ^{222}Rn), equilibrium will be reestablished, that is, in less than 28 days. In practice, it is highly unlikely that all of the radon in the sample was lost during crushing. Thus, equilibrium could be reestablished in five or six half-lives (say 3 weeks), and sample analysis by GRS would be valid (Scott and Dodd, 1960).

Example 3. Radioactive disequilibrium is accepted as the general case in roll-front or sandstone-type uranium deposits (Hansink, 1976). The reason is that uranium is mobile within the sandstone and daughter product formation lags behind. This leads to a distribution of radioelements wherein the daughter products (e.g., ^{214}Bi) are left behind, creating a daughter-excess or parent-deficiency state, with strong gamma ray activity, while at some nearby location, there is a (relatively) weak radioactive uraniferous zone with a daughter deficiency.

For most other rocks, not a great deal is known about the state of radioactive equilibrium in general, although studies of specific areas have been done (e.g., Richardson, 1964). Disequilibrium investigations of the Elliot Lake, Ontario, uranium mining area by Ostrihansky (1976) showed that disequilibrium can occur on a small scale along joints or fractures. Killeen and Carmichael (1976) pointed out that the effect of radioactive disequilibrium on gamma ray measurements is minimized by large sample measurement volumes. Whereas a hand specimen taken from an outcrop might show radioactive disequilibrium, an in situ assay by portable gamma ray spectrometer on the same outcrop comprises such a large sample volume that it may be effectively in equilibrium, that is, the parents and daughters may have moved apart on the scale of a hand specimen, but not on the scale of a cubic meter of rock. Similarly, both parent and daughter nuclides, even if separated on a small scale, could be included in the large surface area and sample volume 'seen' by an airborne gamma ray spectrometer (AGRS), making it more likely that the equilibrium assumption is valid.

11.14.2.4.1 Radon measurement as a radiometric technique

Although radon loss and the consequent radioactive disequilibrium may have a negative effect on the validity of gamma ray spectrometric analyses, in fact, the measurement of radon gas emanation forms the basis of some radiometric techniques used in uranium exploration. Several methods have been developed, which detect the alpha particles emitted by the decaying radon. One method (using the emanometer) involves pumping soil gas from a probe inserted into the ground, into a chamber where the alphas can be counted. A radon determination apparatus was developed and described by Dyck (1969), which is still used in uranium exploration today. Other methods involve burying a detector in a shallow hole for several days or weeks, thus averaging the radon variations and alpha counts. The track-etch method uses an alpha-detecting film (nuclear emulsion) inside a buried inverted plastic cup. When the cup is recovered, the film is etched to enhance the tracks made by the alpha particles, which are then counted under a microscope. Another version uses an electronic 'cup,' which is also buried, but which has a semiconductor alpha particle detector and electronics to record the counts, which are later read by another instrument when the cup is retrieved. It can be reset and reused. An overview of these radon measurement techniques was given by Morse (1977). Tanner (1964, 1980) reviewed radon migration in the ground and Morse (1976) described radon measurements in uranium exploration. The radon measurement techniques are also used to monitor radon in the environment as it relates to public health concerns.

11.14.2.4.2 Equivalent uranium and equivalent thorium

Gamma ray spectrometric determinations of uranium and thorium are based on measurement of gamma radiation from the decay of ^{214}Bi in the uranium (^{238}U) decay series and from ^{208}Tl in the thorium (^{232}Th) series, assuming secular

equilibrium. Consequently, what is determined is the concentration of uranium or thorium equivalent to what would be in equilibrium with the measured radioactivity of the bismuth or thallium isotope. The results of these gamma ray analyses are referred to as equivalent uranium (eU) and equivalent thorium (eTh). The IAEA (1976) recommended that the indirect determinations of U and Th concentrations by GRS be denoted by a prefixed 'e' (e.g., eU and eTh, and the U/Th ratio becomes the eU/eTh ratio). In the case of potassium, the radioactivity of the potassium isotope (^{40}K) is measured directly.

11.14.2.5 The Theoretical Gamma Ray Energy Spectrum

The gamma rays emitted by the daughter nuclides of uranium and thorium were tabulated by Smith and Wollenberg (1972). A summary of their table of gamma ray energies and intensities is given by Grasty (1979), and a more recent tabulation is in IAEA (2003). The principle gamma rays of the uranium and thorium series have also been tabulated in an appendix of Adams and Lowder (1964). Most compilations like these are derived from Lederer and Shirley (1978), Lederer et al. (1968), or Hydeet al. (1964).

The theoretical energy spectra in **Figure 2** (IAEA, 2003; Killeen, 1979) show the principal gamma rays emitted by the ^{238}U decay series (top) and by the ^{232}Th decay series (bottom). The potassium (^{40}K) spectrum (not shown) is a single gamma ray peak at 1.46 MeV. The relative activities or intensities are those obtained under conditions of radioactive equilibrium. Disequilibrium may introduce differences in the relative peak heights depending on whether any given daughter emitter has been removed or deposited in the sample.

However, due to the inherently imperfect energy measuring capabilities of scintillation detectors and scattering of gamma rays and degradation of gamma ray energies in the source rock or soil, the measured spectra for the uranium and thorium decay series will have broad peaks and valleys instead of the narrow spikelike peaks of the theoretical energy spectra. The measured spectra are discussed later.

11.14.2.6 Interaction of Gamma Rays with Geologic Materials

There are three types of interactions between natural gamma rays and materials of concern in gamma ray measurements for geophysical applications, including geologic materials, water, and air. These interactions, which cause attenuation of gamma rays, are Compton scattering, the photoelectric effect, and pair production. In Compton scattering, a gamma ray photon collides with an electron and some of its energy is imparted to the electron. The photon, now with lower energy, is scattered at an angle to its original direction. Compton scattering is the dominant interaction at moderate gamma ray energies in most common geologic materials. After numerous collisions and loss of energy, the dominant interaction for low-energy gamma rays becomes the photoelectric effect, in which the photon loses all of its remaining energy in ejecting a bound electron from an atom. The probability of a photoelectric interaction also increases with increasing atomic number Z of the materials. Rocks are a complex medium containing many different elements, and the atomic number of rocks is referred to as the 'equivalent' atomic

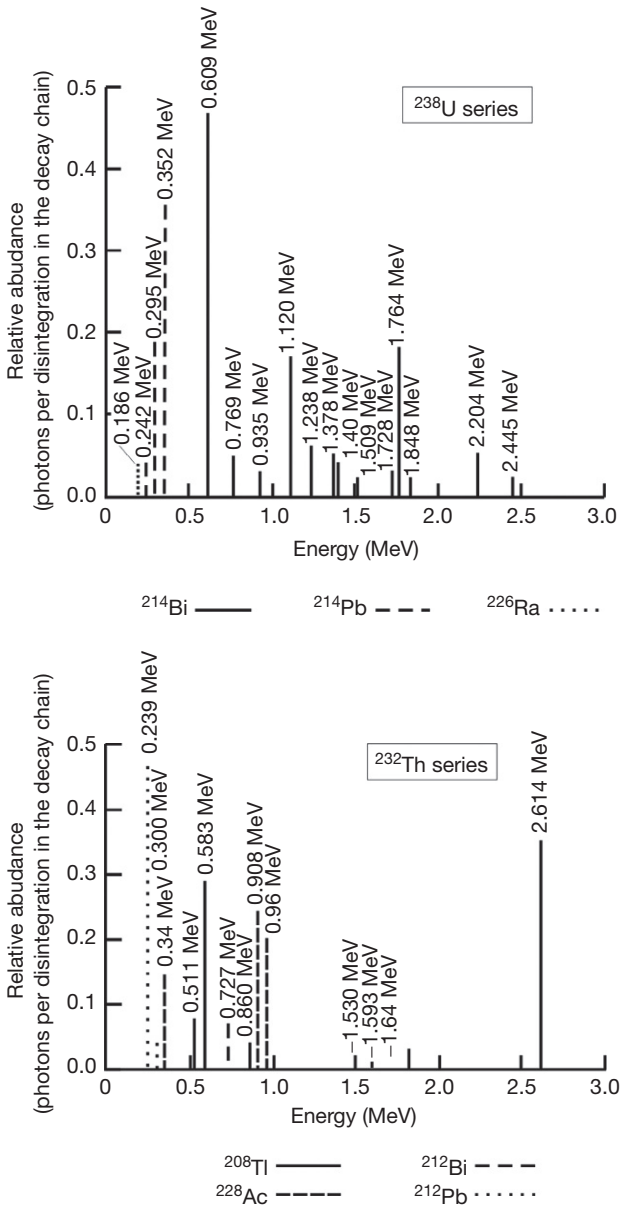


Figure 2 Theoretical natural gamma ray spectra showing the energies and relative intensities of the principal gamma rays emitted by the ^{238}U (top) and ^{232}Th (bottom) radioactive decay series. Reproduced from Killeen PG (1979) Gamma ray spectrometric methods in uranium exploration: Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 163–230. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

number Z_{eq}. The Z_{eq} of a complex medium is defined (IAEA, 1982) as the atomic number of the element having the same ratio of Compton attenuation to photoelectric attenuation as the complex medium (for a given photon energy).

Pair production only occurs for gamma rays with energies < 1.02 MeV when the photon is completely absorbed and the energy produces an electron–positron pair. The probability of pair production also increases with increasing Z. Pair production is not significant in most geologic materials, which have a Z_{eq} ranging from about 12 to 18. However, in gamma

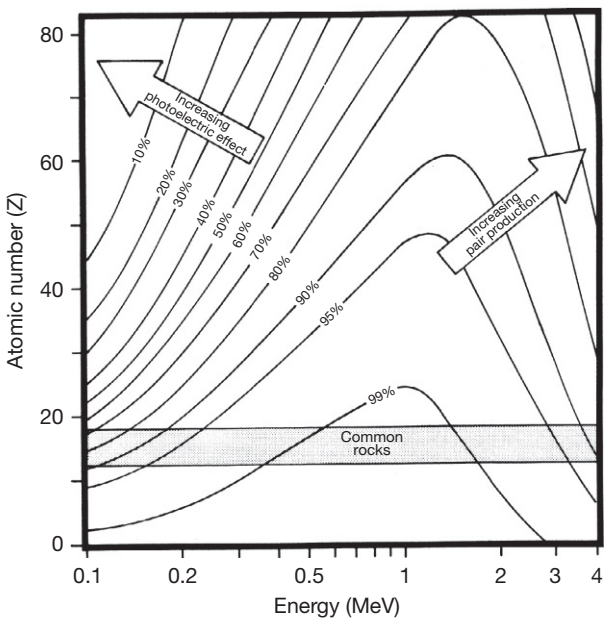


Figure 3 The proportion of the three gamma ray interactions for different energies and materials of different atomic number Z. The percentage of Compton scattering interactions is contoured, showing it is the dominant interaction for the range of Z_{eq} for common rocks (about 12–18) and for the range of natural gamma ray energies (about 0–3.0 MeV). Reproduced from Conaway JG (1980) Uranium concentrations and the system response function in gamma ray logging. In: *Current Research, Part A*, pp. 77–87. Ottawa, ON: Geological Survey of Canada, Paper 80-1A.

ray logging to evaluate uranium deposits, the high Z (92) of the uranium in anomalous concentrations causes the so-called Z-effect in which the uranium strongly attenuates its own gamma rays primarily by the photoelectric effect.

The proportion in which these three interactions occur depends on gamma ray energy and Z_{eq} as shown in Figure 3. For the range of atomic numbers in common rocks ($Z = 12$–18), and the energies of natural gamma rays (from about 0 to almost 3 MeV), about 95% of the interactions are Compton scattering. Because of these three interactions, the range of penetration of gamma rays in geologic materials is limited by the gamma ray energy, the density (which is related to the electron density), and the Z_{eq}. The intensity of gamma radiation from a natural radioactive source decreases exponentially with distance from the source. This has ramifications for the determination of the sample volume of any gamma ray measurement (to be discussed later). For common rocks, the penetration range for natural radiation is approximately 0.5 m. Knowledge of the sample volume is important in quantitative measurements and for planning the physical dimensions of calibration facilities.

11.14.2.7 Detecting and Counting Gamma Rays

11.14.2.7.1 Geiger–Müller counters

Although the Geiger counter originated earlier as a heavy instrument in the lab, shortly after 1934, Ellsworth at the Geological Survey of Canada (GSC) built a 'portable' (9.5 kg) version suitable for use in field measurements (Haycock, 1953). Radiometric techniques in Earth sciences (outside the lab) probably first started with the search for radioactive minerals around that time, but exploration for uranium deposits after 1943 greatly expanded the use of Geiger counters. Geiger counters detected gamma rays by the ionization, which they produced in a gas-filled tube. Operating at hundreds of volts between an anode and cathode, an electric field swept out the ionization to produce a current and voltage pulse that could be counted. The detection efficiency (ratio of counts recorded to the number of gamma rays that struck the detector) was about 2%, which is extremely low. There was no capability to distinguish differences in energies of gamma rays and only total count (TC) rate was measured. The dead time (time to process a pulse, during which any other gamma rays (pulses) are not able to be counted) was also considered high at about 0.1 ms. One application where Geiger tubes are still in use today is in borehole gamma ray logging in holes penetrating uranium deposits. There, the gamma radiation level is so high that the inefficient detector is an advantage, avoiding the saturation (overloading) effects in more efficient detectors.

11.14.2.7.2 Scintillometers and scintillation detectors

A scintillometer is basically an instrument with a scintillation detector composed of a crystal and photomultiplier tube (PMT), a high-voltage supply, an amplifier, and a counter. It can count gamma rays (i.e., detect their presence) but measure only the total radioactivity. A scintillometer can be used to locate anomalous sources of radioactivity.

A scintillation detector operates on the principle that gamma rays produce scintillations or flashes of light in a scintillation crystal such as sodium iodide, which are 'seen' by a PMT in contact with the crystal. To make the crystal produce scintillations, small amounts of 'activators' are introduced into the crystal such as in thallium-activated (or 'thallium-doped') sodium iodide (NaI(Tl)), which is grown from a melt. The output of the PMT is a voltage pulse with amplitude proportional to the energy of the gamma ray, making them suitable for use in gamma ray spectrometers (Figure 4). Scintillation detectors are very efficient (up to 100%) and have dead times that are a thousand times less than that of Geiger detectors. There are numerous scintillation materials available, but the most widely used in geophysical radiometric techniques is thallium-activated sodium iodide (NaI (Tl)) with a density of 3.67 g cm^{-3}. The crystals are usually (but not always) in a cylindrical shape with the PMT attached to a 'window' on the end of the cylinder, which is hermetically sealed because NaI is hygroscopic. A high-voltage supply is used for the PMT and the output voltage pulses are amplified and sent to a counter in the case of a gamma ray scintillometer or to a pulse height analyzer (PHA) in the case of a gamma ray spectrometer. To prevent the scintillometer from counting low-voltage noise pulses as gamma rays, some method of rejecting pulses below a certain level is implemented in the instrument. This effectively also determines the lowest-energy gamma ray that can be counted. Thus, energy calibration, as well as dead time corrections, is required, as described in the succeeding text for spectrometers.

The dimensions of scintillation crystals are always given as the diameter of the cylinder multiplied by the length, for

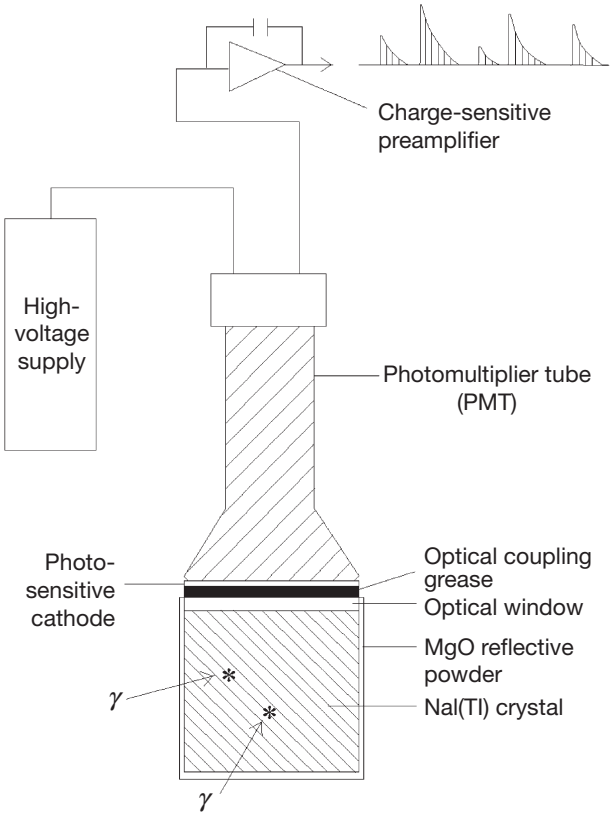


Figure 4 Scintillation detector components showing gamma rays striking the scintillation crystal attached to the photomultiplier tube (PMT), the high-voltage supply, and the preamplifier. The output voltage pulse produced by each gamma ray has an amplitude, which is proportional to the gamma ray energy. Reproduced from Bristow Q (1979) Gamma ray spectrometric methods in uranium exploration: Airborne instrumentation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 135–146. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

example, a 25×10 detector has a cylindrical crystal with a diameter of 25 cm and a length of 10 cm. The exception is the modern large crystals used in AGRS, which are square in cross section and relatively long. Typical dimensions are $10\ \text{cm} \times 10\ \text{cm} \times 41\ \text{cm}$ ($4 \times 4 \times 16$ in.). The PMT attached to the square end has a circular shape just as when attached to cylindrical crystals.

11.14.2.7.2.1 Energy resolution of the detector

If a detector is used to measure the energy of gamma rays as in a gamma ray spectrometer, the detector's energy resolution is very important. The 'resolution' of a scintillation detector such as a NaI (Tl) crystal is a measure of its ability to distinguish the difference in energy of gamma rays with similar energies. In other words, if the detector 'sees' the gamma ray energy spectrum as a number of narrow sharp peaks, it has good resolution, and if the peaks are broad and flattened, the detector has poor resolution. The international standard agreed upon to determine the energy resolution of detectors is based on measuring the gamma ray spectrum of the 662 keV gamma ray peak of ^{137}Cs (IAEA, 1991). The energy resolution is defined as the ratio of the full width (in keV) of the peak measured at half the maximum peak height (FWHM), expressed as a percentage of the energy of the peak (662 keV). For example, a FWHM of 66.2 keV equals a resolution of 10%. Usually, the resolution ranges from about 6% (extremely good) to 20% (extremely poor). Most detectors used in airborne gamma ray spectrometers have energy resolutions ranging from about 7% to 10%. A lower value is better because it results in sharper peaks in a measured gamma ray energy spectrum. Resolution varies with detector size as well as detector material. Different scintillation crystals of the same material, for example, NaI (Tl), can have a range of resolutions, and it must be measured for each detector. The cost of detectors varies with their energy resolution, the best resolution detectors for use in exploration commanding higher prices. (Note: The energy resolution may be measured on other peak energies than 662 keV, but the result must be converted to the equivalent of measurement on the 662 keV peak using the fact that the resolution is inversely proportional to the square root of energy (Knoll, 1989).)

11.14.2.7.3 Gamma ray spectrometers

A gamma ray spectrometer, like a scintillometer, has a scintillation detector (crystal and PMT), a high-voltage supply, and an amplifier to detect gamma rays, but it also has a PHA, which can measure the height (amplitude) of the voltage pulses from the PMT before they are counted. The pulse height is proportional to the energy of the gamma ray that caused it. Thus, they are a measure of the gamma ray energies, and the pulses can be sorted into a histogram of energies of gamma rays, or an energy spectrum. Since the gamma ray energies are specific for each radioactive element, a gamma ray spectrometer can be used to identify whether the source is predominately potassium, uranium, thorium, or a combination, based on their peaks in the energy spectrum.

Early spectrometers used threshold discriminators, electronic devices that would allow voltage pulses above a preset threshold to be counted and reject those below the threshold (or vice versa). Thus, all gamma rays above a preset energy could be counted. With two discriminators, a lower-level discriminator (LLD) and an upper-level discriminator, called a single-channel analyzer (SCA), it was possible to count gamma rays in a window (channel) over a range of energies in the spectrum. Early four-channel 'threshold' spectrometers had one LLD set at about 400 keV to provide the TC and three LLDs set just below the three main gamma ray peaks of the K, U, and Th spectra at 1.46, 1.76, and 2.62 MeV, respectively. 'Differential' four-channel spectrometers had one LLD set at about 400 keV for TC and, in addition, three SCAs with energy window set to straddle the peaks at 1.46, 1.76, and 2.62 MeV to count gamma rays in those peaks from K, U, and Th, respectively. These have evolved into multichannel spectrometers with 256 or more, narrow (energy) channels, covering the full gamma ray energy spectrum. A window is then the sum of counts in a range of channels.

11.14.2.7.3.1 Energy calibration

When making measurements with a gamma ray spectrometer, the relationship between pulse heights (voltage) and gamma ray energies must be established. This is called 'energy calibration.' The amplitude of the pulses is adjusted with an amplifier (i.e., adjusting the gain) until the pulses are counted

in the desired energy window, which is really a range of pulse amplitudes or voltages. Initial calibration is done with two calibration sources such as ^{137}Cs (which emits 662 keV gamma rays) and ^{60}Co (with gamma rays of 1.17 and 1.33 MeV), which are used to produce a straight line graph of gamma ray energy (vertical axis) versus pulse height, that is, a linear relationship over the range of energies of interest. (Details on linearity are discussed further in Section 11.14.3.3.4.) The slope of the line is in units of keV V^{-1} and is also the 'gain.' Thus, any pulse height can then be converted to gamma ray energy. Once this is established, voltage ranges, or energy windows (also called regions of interest), can be set to straddle any gamma ray peaks of interest. The energy calibration of the spectrometer can be checked periodically with a single calibration source.

Although the previous assumptions regarding linearity are widely used over a large range of energies, they are not strictly valid. The use of a quadratic equation for the relation between window (channel) position and energy was described by Grasty et al. (2001).

11.14.2.7.3.2 'Drift' and stabilization

An energy-calibrated spectrometer may drift out of calibration if the gain changes for any reason. Some of the electronic components are temperature-sensitive and the pulse height for a given energy gamma ray may change (i.e., a gain change) such that it is not counted in the proper window. The output of the pulses from the PMT is very temperature-sensitive, for example. Several different methods of gain stabilization have been developed, which are described in Section 11.14.3. As technology advanced, spectrometers began to use analog to digital converters (ADCs) to digitize the pulses and sort the counts into a digital spectrum of gamma ray energies in a multichannel analyzer of say128–256 channels. Modern AGRS commonly record spectra of 512–1024 channels. Energy calibration (in this case with the gain measured as keV/channel) and stabilization are still part of the requirements.

11.14.2.7.3.3 Dead time

Another factor to be considered in any gamma ray counting is dead time. The recorded counts must be corrected to account for dead time and there are several methods to do this described in Section 11.14.3. Since it takes a finite amount of time to detect a gamma ray, measure its energy, and count it, the instrument will be 'dead' while it is busy. If any gamma rays strike the detector during the dead time, they will not be counted. In practice, in many modern instruments, the dead time is so small it can be ignored or corrected automatically in the electronics. If the dead time (usually in μs) per pulse (gamma ray recorded) is known for an instrument, there is a simple equation to obtain the true count rate from the observed count rate. This will be elaborated upon in Section 11.14.3.

11.14.2.7.3.4 Background radiation

Background radiation is any radiation detected by the gamma ray spectrometer not originating from the source that is being analyzed, in this case the lithosphere (IAEA, 1976). In the case of a laboratory spectrometer, it includes radiation coming from or through the walls, ceiling and floor, and the lead counting chamber or shield. In the field, it includes the radiation from the vehicle, be it a man carrying the spectrometer, a truck, or an aircraft. In addition, there is a cosmic ray component (most important for airborne surveys) and radioactivity in the atmosphere caused by radon and its daughter products and products from nuclear fallout. To the extent possible, the background should be minimized. In an airborne survey, for example, any flight instrument dials or emergency exit signs, etc. in the aircraft, which are luminized with radium (found in older aircraft), should be replaced or removed. Calibration sources (as described in Section 11.14.2.7.3.1) should be removed completely or at least be shielded if they must be carried on the survey. Any remaining background should be measured accurately to enable its subtraction. The greatest problem with background is the variable or unknown component primarily caused by the radioactivity of particles in the air. Background measurement was reviewed in some detail especially with reference to airborne gamma ray surveys (Grasty, 1979; Grasty et al., 1988; IAEA, 2003). Generally, the background radiation in a borehole gamma ray log is considered to be zero except in a few exceptional cases, since the detector is surrounded by the source. Cosmic ray effects are generally negligible, and the radioactivity of the air has no influence in a liquid-filled hole and negligible influence in an air-filled hole (except in uranium exploration holes) due to the small volume of air.

11.14.2.8 The Measured Gamma Ray Energy Spectrum

The theoretical natural gamma ray energy spectrum was discussed in the preceding text (see Figure 2). However, the measured gamma ray energy spectrum is smeared, and the discrete well-separated energy peaks show as rounded peaks, which overlap like those shown in Figure 5. The amount of broadening of the peaks and filling of the valleys between them depends partly on the energy resolution of the detector and partly on the Compton scattering interactions of the gamma rays with the geologic materials before striking the detector. Depending on the number of collisions with electrons (Compton scatter), gamma rays might arrive at the detector with any energy ranging from its initial energy to almost zero energy and thus get counted in lower energy channels of the spectrum. This gives rise to the 'Compton continuum': the spectrum of all the gamma rays that are reduced in energy, which are counted in a lower energy channel. This makes the measured spectral peaks appear to be riding on a sloped spectrum with high counts at low energies sloping down to low counts at high energies. In addition to the previously mentioned two factors, the gamma ray may not be completely stopped by the detector, and the scintillations will produce a pulse equivalent to a gamma ray with lower energy than the true energy.

The energy spectra measured with a NaI (Tl) detector, of gamma rays emitted by pure sources of ^{40}K (top), the ^{238}U decay series (middle), and the ^{232}Th decay series (bottom), are shown in Figure 5. The IAEA-recommended and commonly used windows for natural GRS are indicated in the figure. The windows straddle the three most important peaks for K, U, and Th decay series, which are at 1.46 MeV (^{40}K), 1.76 MeV (^{214}Bi), and 2.62 MeV (^{208}Tl), with window widths of 200, 200, and 400 keV, respectively. Table 5 is a list of the TC,

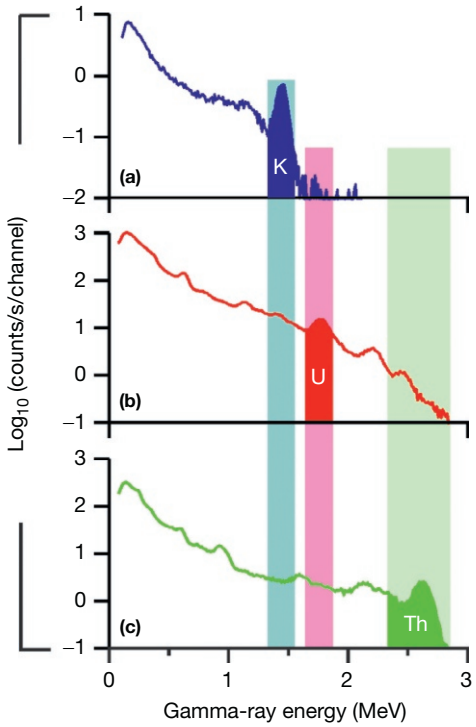


Figure 5 Measured gamma ray spectra: (a) potassium, (b) uranium, and (c) thorium, from K, U, and Th model borehole calibration facilities. The three standard energy windows recommended by the IAEA for gamma ray spectrometric measurements of the natural radioelements are shown. They straddle the 1.46, 1.76, and 2.62 MeV gamma ray peaks used to analyze for K, U, and Th, respectively. It is easy to see the interference in the windows among the three spectra. For example, the thorium spectrum will contribute gamma counts to the U and K windows as well as to the Th window. These unwanted counts must be stripped from the windows using a set of stripping factors.

potassium, uranium, and thorium windows, the element, which is the gamma emitter, the energy of the peak, and the energy range for each window.

Most portable gamma ray spectrometers recorded only the counts in the three windows, not the entire spectrum. In addition, a TC window recorded all gamma rays greater than a certain energy, usually 0.4 MeV. However, today, most modern portable spectrometers also record spectra with similar ranges to those recorded with car-borne and airborne systems. Although large vehicle-borne gamma ray spectrometer and AGRS systems record the entire energy spectrum in 256–1024 channels, most data processing techniques sum the counts in the four 'standard' windows of the multichannel spectra and process the windowed data in much the same way as processed for a portable spectrometer.

In any case, it can be seen in Figure 5 that there is interference between the three spectra, which results in counts being recorded in the wrong windows. For example, it is clear from the figure that some gamma rays originating from the Th decay series will be counted in the U and K windows. These and other interferences between the window counts are corrected using a process called spectral stripping (or Compton stripping).

11.14.2.8.1 Stripping factors

The determination and various definitions of the stripping factors (often called stripping ratios) have been discussed in numerous papers (Adams and Fryer, 1964; Adams and Gasparini, 1970; Doig, 1968; Grasty, 1976b, 1977, 1979; Grasty and Darnley, 1971; IAEA, 1991, 2003; Killeen and Cameron, 1977; Killeen and Carmichael, 1970; Wollenberg and Smith, 1964).

Based on the interference in the windows between K, U, and Th spectra seen in Figure 5 and described above, we can define the stripping factors or stripping ratios. The following somewhat simplified explanation and example clarifies the meaning and application of the stripping ratios. It is equally valid for window counts or count rates in counts per second (cps).

By looking at the thorium spectrum from a pure monoelemental thorium calibration source (Figure 5), we see that the U and K windows have counts from gamma rays originating from the Th decay series. The counts recorded in the Th window (Thc) originate from the thorium decay series. The counts recorded in the U window (Uc) are also from the thorium decay series. For this thorium spectrum, the ratio of the uranium window counts (Uc) (from thorium) to the thorium window counts (Thc) (from thorium) is the stripping factor used to correct the counts in the U window, by stripping off the thorium counts that do not belong there. This stripping ratio is called α:

$$\alpha = \mathrm{Uc}\,(\mathrm{Thc})^{-1} \qquad [5]$$

which is the number of erroneous U counts (from Th), in the U window per Th count in the Th window. Thus, the corrected (for the Th contribution only) stripped count (Us) in the U window is

$$\mathrm{Us} = \mathrm{Uc} - \alpha\ \mathrm{Thc} \qquad [6]$$

Similarly, the K window includes counts from gamma rays originating from the Th decay series, and this is corrected with a similarly derived stripping ratio called β. In addition, using a uranium source, it can be seen that some gamma rays originating from the U decay series will be counted in the lower-energy K window and to a small extent in the higher-energy Th window (requiring an 'upward' stripping ratio), and finally using a K source, we see that some small portion of the K gamma rays may be counted in the U window or even (theoretically) in the Th window. The aforementioned simplified explanation and derivation is based on spectra from pure calibration sources of K, U, and Th. Pure monoelemental calibration sources are difficult to obtain because they are made from naturally occurring K, U, or Th minerals and the three radioelements are usually found together in most geologic materials. Thus, the derivation of the stripping ratios is slightly more complex, requiring a regression analysis to take into consideration the small concentrations of the other two radioelements in each of the three 'pure' sources (Grasty, 1979).

These considerations lead to derivation of six stripping ratios as shown in Figure 6. The figure is a schematic representation of the interaction between the K, U, and Th energy windows, which indicates the stripping ratios, which are used to remove the interference denoted by the arrows. The six stripping ratios are referred by the Greek letters α, β, and γ,

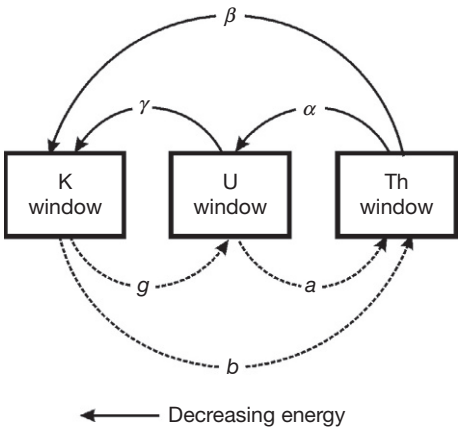


Figure 6 Representation of the six stripping ratios used to correct for the interference between counts recorded in the K, U, and Th energy windows as shown in **Figure 5**. For example, gamma rays in the U window from the Th decay series can be removed by subtracting α multiplied by the Th window counts from the U window count. Similarly, gamma rays counted in the Th window from the U decay series can be removed by subtracting '*a*' times the U window counts from the Th window count. The 'upward' stripping ratios *a*, *b*, and *g* are usually small. Reproduced from Killeen PG (1979) Gamma ray spectrometric methods in uranium exploration: Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 163–230. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

and the 'upward' or 'reverse' stripping factors by *a*, *b*, and *g* (Killeen, 1979).

In summary, the interplay between the three radioelement windows K, U, and Th and the six associated stripping ratios and their purpose is listed in the succeeding text:

- α is used to strip off Th gamma rays in U window.
- β is used to strip off Th gamma rays in K window.
- γ is used to strip off U gamma rays in K window.
- *a* is used to strip off U gamma rays in Th window (usually small).
- *b* is used to strip off K gamma rays in Th window (approximately zero).
- *g* is used to strip off K gamma rays in U window (approximately zero).

For some purposes, only the first three stripping factors are used to simplify computations and *a*, *b*, and *g* are assumed to be zero. A simplified method for computing K, U, and Th concentrations assuming *a*, *b*, and *g* are zero is described by Killeen and Cameron (1977) and, assuming *b*, and *g* are zero, is described in IAEA (2003). For high uranium concentrations, the upward stripping factor '*a*' may be quite significant. It has a value of approximately 0.05 (Grasty, 1975), but the actual value is dependent on factors such as detector size and resolution, window widths, and energy settings. After stripping and other appropriate corrections (e.g., for background) are made, the corrected counts in the windows are ready to be converted to radioelement concentrations via the sensitivities or '*K*' factors. The stripping factors are determined by making measurements on 'nearly' monoelemental concrete calibration pads. The derivation and use of the *K* factor and stripping factors are described in greater detail in the section on calibration. (Note: the *K* factor is often called *C* factor or *S* factor.)

11.14.2.9 Source–Detector Geometry

Since the amount of radiation that can be detected is related to the size and shape of the radiometric source as well as its intensity, the so-called geometry must be taken into consideration. Generally, geometry refers to the angular measurement or solid angle (in steradians), which the source subtends at the detector, where 4π steradians is equivalent to complete enclosure of the detector by the source. For example, in a laboratory, the radiation may be collimated by lead bricks or other absorbers such that the source only subtends an angle of a few degrees at the detector. A detector placed above an 'infinite' planar source would be an example of 2π geometry, and a detector inside a borehole would be in 4π geometry, being virtually completely surrounded by the source.

A gamma ray spectrometer can be calibrated to analyze for the concentration of a radioelement by determining the relationship between the measured gamma ray count rate and the known radioelement concentrations of calibration sources. The calibration of any gamma ray spectrometer is for a specific geometry. For this reason, it is important to know the geometry of the intended field measurements. For example, because of the difference in geometry, a hand specimen of 1% U_30_8 ore will produce a much lower count rate than a rock outcrop averaging the same 1% U_30_8, even if the detector is located at the same distance from the source. To make sense of recorded count rates, the geometry of the measurement must be noted. A few examples are given later to illustrate situations in which the geometry of a measurement can change to produce a false anomaly or to mask a real anomaly.

11.14.2.9.1 Borehole measurement geometry

In the case of a gamma ray probe inside a borehole, the source effectively surrounds the detector except for a narrow solid angle above and below it. In an air-filled borehole, the geometry will still be 4π even if a large cave-in or washout is encountered, because the rock surrounds the detector. However, in a water-filled hole, the attenuation of the additional water in the enlarged portion of the hole requires corrections to be made.

11.14.2.9.2 Surface measurement geometry

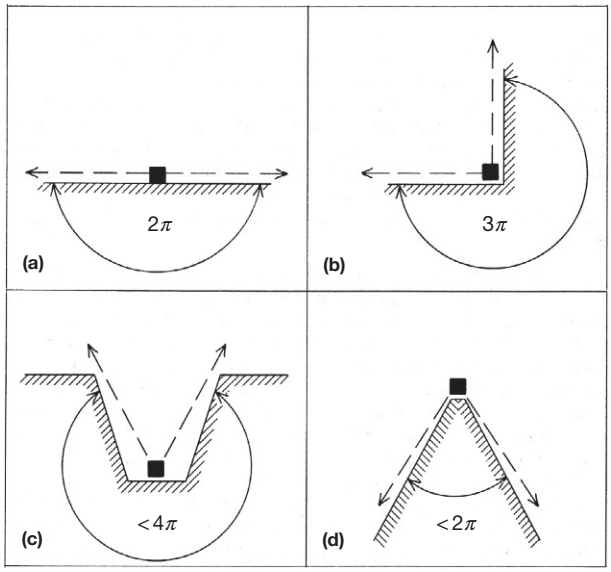
Generally, radiometric measurements made above the surface of the Earth by either handheld portable gamma ray detectors, vehicle-mounted detectors, or airborne detectors are considered to be measurements in a 2π geometry. The instruments should be calibrated on flat calibration sources of effectively infinite diameter, such as the concrete pads discussed in the section on calibration facilities. In the field, on the ground, a measurement made in a trench or near a cliff face effectively changes the geometry to $>2\pi$ such that use of the 2π calibration factors would be invalid. A change in the geometry can occur during a car-borne survey, for example, if the vehicle passes through a roadcut creating a false anomaly.

Some early work described the use of lead shielding to control the geometry in the measurement of radioelement concentrations with a portable gamma ray spectrometer (Lovborg et al., 1969; Mahdavi, 1964). Corrections for the

effects of geometry on measurements with a portable gamma ray spectrometer were reported by Wormold and Clayton (1976). More recently, work on a directional spectrometer based on the use of multiple detectors and coincidence counting methods was reported (Khan, 2012). A schematic of four different source–detector geometries encountered in surface gamma ray spectrometric measurements in the field is shown in Figure 7.

11.14.2.9.3 Airborne measurement geometry

There is not much variation from 2π geometry for a fixed-wing airborne survey if the terrain is not too rugged. Large fixed-wing aircraft carrying large detector arrays for reconnaissance work are less maneuverable than the smaller aircraft or helicopters usually used in detailed surveys. In rugged terrain, maintaining a constant elevation is difficult and may become impossible with a fixed-wing aircraft. Aircraft elevation corrections can be applied, but geometry corrections can usually be applied only in a qualitative sense or at best semiquantitatively. The effect of source geometry and planned flight height h upon count rate over flat ground, over broad topographic changes, over sharp hills, and in narrow valleys was discussed by Grasty (1976a). In these last two cases, the geometric configuration is $<2\pi$ over hills and $>2\pi$ in narrow valleys. The height will be less than h and more than h, respectively, depending on the geometry. The height correction factors can compensate for this, but not for the changes in geometry.



(a) 2π geometry
(b) 3π geometry
(c) $< 4\pi$ geometry
(d) $< 2\pi$ geometry

Figure 7 Four possible source–detector geometries seen in surface gamma ray measurements, (a) on a flat surface, (b) near a cliff face, (c) in a trench, and (d) on a hill top. Reproduced from Killeen PG (1979) Gamma ray spectrometric methods in uranium exploration: Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 163–230. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

11.14.2.10 Effective Sample Volume

The size of the sample being analyzed in a field measurement has been discussed by several authors. Lovborg et al. (1971) did considerable work with respect to portable gamma ray spectrometers, Grasty (1979) considered the sample volume of airborne radiometric measurements, and Dodd and Eschliman (1972) considered the borehole measurement case for TC logging. Several factors can affect the sample volume, such as the energy of gamma radiation being measured, the density of the source material (rock, overburden, etc.), the absorption coefficient of the material, and whether the detector is moving or stationary. The definition of sample volume is also important. For example, it could be arbitrarily defined as the volume within which 90% of the detected gamma rays originated.

The difficulty of defining the sample volume when detecting gamma rays of different energies in materials of different densities was illustrated with graphs of attenuation (intensity/initial intensity, I/I_0) versus thickness in Killeen (1979). The curves were for three gamma ray energies (1.46, 1.76, and 2.62 MeV for K, U, and Th, respectively) and three rock densities (2.0, 2.67, and 3.0 g cm^{-3}, for low, average, and high). The resulting curves showed that for the low density of 2.0 g cm^{-3}, attenuation of nearly 100% occurred at about 50 cm for 2.62 MeV gamma rays and at 40 cm for 1.46 MeV gammas, while for the high density of 3.0 g cm^{-3}, 100% attenuation was at 40 cm for 2.62 MeV gammas and at 30 cm for 1.46 MeV gamma rays. The conclusion is that the penetration range of natural gamma rays in geologic materials varies from about 30 to 50 cm, making it difficult to define the sample volume. The sample volume is also dependent on the area viewed by the detector.

11.14.2.10.1 Borehole measurement sample volume

To illustrate the difficulty of defining the boundaries of the sample 'volume' for a TC gamma ray log, Figure 8 shows a sketch of a vertical cross section along the axis of a borehole through an infinitely thick homogeneous zone of radioelements. A scintillation detector in a probe is shown in the borehole with an array of dots in the rock surrounding the detector. The dots represent the distribution of radioactive sources, which emitted the gamma rays detected during a sample interval. The density of the dots (sources) decreases exponentially away from the detector, showing that it is virtually impossible to include 100% of the sources in the definition of the sample volume. In addition, the volume would include a cylinder of length L if the detector moves a distance L during the time of measurement.

The sample volume for a borehole gamma ray detector was illustrated as a series of concentric circles or annular zones surrounding the detector in the hole, by Dodd and Eschliman (1972). The circles were located at a radial distance from the detector marking the percentage of detected gamma rays that originated from inside the circle. It was determined that 50% of the detected gamma rays came from <6.6 cm outside the hole and 90% came from <18 cm. They considered that an annular zone of radius 40 cm was infinitely thick, so the infinite sample volume would be about 80 cm diameter. Computations were based on a density of 2.2 g cm^{-3}.

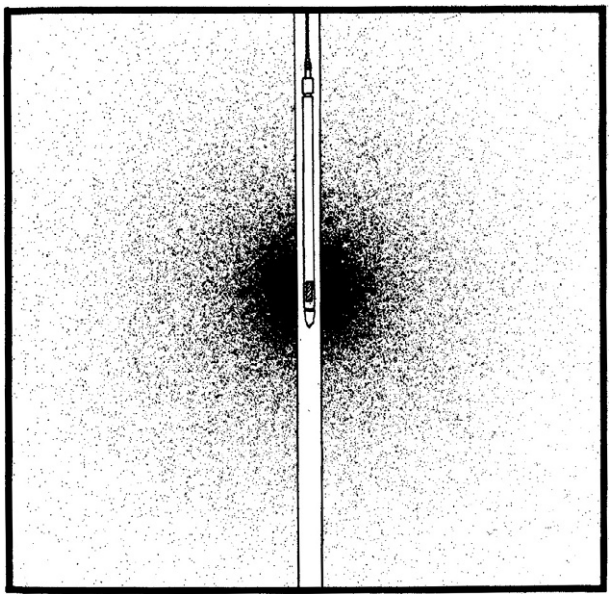


Figure 8 Sample volume for a borehole gamma ray detector, shown as dots representing the distribution of radioactive sources, which emitted the gamma rays detected during a sample interval. The sample volume is often defined as the region of rock around the detector from which a given percentage (e.g., 90%) of the gamma rays striking the detector originated. Reproduced from Killeen PG (1979) Gamma ray spectrometric methods in uranium exploration: Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 163–230. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

11.14.2.10.2 Surface measurement sample volume

The sample volume for an in situ assay by a portable gamma ray spectrometer would essentially be half that shown in the borehole 4π case in Figure 8. The sample volume was illustrated by Lovborg et al. (1969) as a vertical cross section with a detector on the ground surface, and below the surface, contour lines represented nearly hemispherical shells each contributing 10% of the detected gamma radiation for the case of the 2.62 MeV gamma rays from the ^{232}Th decay series. They showed that about 90% of the gamma rays detected originated from a depth of about 35 cm and a horizontal radial distance from the detector of roughly 75 cm. Values were computed using a density of 2.85 g cm^{-3}.

While studying the design of calibration pads for calibrating portable gamma ray spectrometers, Grasty et al. (1991) determined the percentage of an infinite source represented by 50 cm thick pads with different diameters for a detector positioned 6 cm above the pad surface. For the gamma rays in the thorium window at 2.62 MeV, a pad diameter of 50 cm represented 60% of an infinite source, 100 cm represented 83% of an infinite source, and a 3 m diameter pad was 95.6%.

11.14.2.10.3 Airborne measurement sample volume

In an airborne gamma ray spectrometric survey, the volume of the sample is the product of the surface area 'seen' by the airborne detector and the thickness of the source material. The sample volume for an airborne measurement will depend to a large extent on the height of the aircraft above the ground

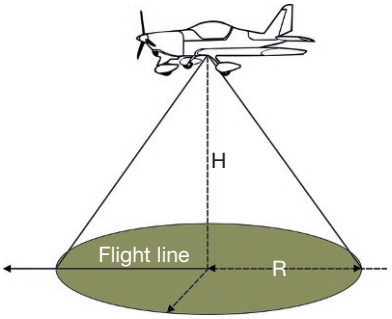


Figure 9 The concept of the field of view for an airborne detector, which relates to the sample volume. The radius of the circle of investigation varies with altitude and the percentage of an infinite source that it is meant to represent. Also, during a given sample interval, the circle moves along the ground stretching it into a quasi-ellipse of investigation.

(Figure 9). This affects the area of the surface that is analyzed. The area of the surface has been called the field of view (FOV), instantaneous FOV, circle of investigation, and area of influence by various authors.

Calculations of the size of the sample area viewed by an AGRS were presented by Duval et al. (1971) and by Grasty (1976a,b) and Grasty et al. (1979). The computed radius of the circle of investigation plotted as a function of altitude for a given percentage (from 25% to 95%) of infinite source yield was given by Duval et al. (1971) for a stationary detector. For example, the curves show that at height h of 120 m, 95% of an infinite source of gamma rays had a radius of about 3 h or 360 m. Further work on the same type of computations was presented by Clark et al. (1972) and by Adams and Clark (1972).

Grasty et al. (1979) determined that the width of the area supplying a given percentage of the radiation detected is not the same for measurements made with the aircraft moving or stationary. They calculated the area of coverage for an airborne gamma ray spectrometric survey for a stationary detector and a moving detector at altitudes of 120 and 25 m. For a given percentage of infinite source detected, the width of the strip for a moving detector is less than the diameter of circle for a stationary detector. Calculations of the amount of surface area covered for a given airborne survey, based on the stationary detector model, will produce overestimates, since the survey is actually done with a moving detector. The use of the concept of the 'FOV' in computing the percentage coverage of an area in a given survey is important in choosing the desired flight line spacing for an airborne survey.

The diameter of the circular area beneath the aircraft contributing a given percentage of the gamma radiation varies with detector size, with the diameter being larger for a smaller detector (Grasty, 1979). This is due to the differing sensitivity of the detectors in different directions. The data show the change in diameter can be in the order of 10% variation. It also varies with gamma ray energy; the higher the energy, the larger the diameter of the circle.

The percentage of an infinite source represented by circular, 50 cm thick, calibration pads of different diameters was determined by Grasty (1987) for calibrating AGRS with a detector 1 m above the pad, typical for an aircraft parked on the pad.

A 4 m diameter pad represented 77% of the radiation from an infinite source, and an 8 m diameter pad represented 89%.

11.14.3 Instrumentation

11.14.3.1 Introduction

The radiometric technique most widely used in near-surface geophysics is GRS, which can be used in borehole, surface, and airborne configurations. This is because spectrometry can be used as a quantitative method for radioelement analysis, whereas scintillometers are used for qualitative measurements. Portable gamma ray spectrometers, even handheld versions, can be calibrated not only to assay for naturally occurring radioactive elements (e.g., potassium, uranium, and thorium) if properly calibrated but also to detect and identify man-made radioactive sources such as radioactive contamination, which might be found in environmental surveys after a nuclear incident. The identification is based on measuring the characteristic gamma ray peaks in the energy spectra of the sources.

Gamma ray spectrometric instrumentation is basically the same for borehole, surface, and airborne systems with major differences being related to detector size and type, as well as the degree of sophistication of the data acquisition and signal processing, including the sampling rate. For example, an airborne spectrometer moves very fast and must make measurements in rapid succession (e.g., once per second or faster), whereas a borehole logging system moves relatively slowly and can sample less frequently. Spectrometers have evolved from simple three- or four-window instruments to multichannel versions now used routinely not only in airborne systems but also in surface and borehole gamma ray spectrometers. However, often, even the multichannel spectrometer data are ultimately distilled down to the three-window data for processing in a similar fashion to the simplest earlier instruments.

The following is directly applicable to all gamma ray spectrometers but is focused on AGRS as they are more widely used in comparison to surface and borehole spectrometers.

11.14.3.2 Counting Statistics

An important requirement for a radiometric survey is that it must have 'good counting statistics.' This is a rather broad term but basically it means that the gamma ray count rate or total number of gamma rays counted in a given measurement must be large enough to be considered a statistically reliable measurement. This is related to the desired accuracy of a measurement (Chase and Rabinowitz, 1968; IAEA, 2003).

Radioactive decay is a random process, and one standard deviation (σ) in a counting measurement equals the square root of the count (n) obtained. This can then be expressed as a percentage of the count or percentage error. One standard deviation is

$$\sigma = \sqrt{n} \qquad [7]$$

Thus, the error in a measurement where $n = 100$ counts is ± 10 counts, or a percentage error of $\pm 10\%$. A count of 1000 would yield ± 31.6 counts, about $\pm 3\%$ error.

The standard deviation in the counting rate is

$$\sigma_{\mathrm{RATE}} = \sqrt{nt^{-1}} = \sqrt{\left(\mathrm{RATE}\ t^{-1}\right)} \qquad [8]$$

where the counting rate is defined as $\mathrm{RATE} = nt^{-1}$.

The error in the final result will be increased when a number of corrections are applied such as background subtraction and spectral stripping, each of which includes a counting measurement with its associated errors.

To obtain a lower error, the count must be increased; this can be accomplished by (1) counting for a longer time, (2) increasing the detector size, or (3) moving the detector closer to the source. These three alternatives are relatively easy to accomplish in a laboratory, but much more difficult in field measurements. Increasing the counting time for a portable (man-carried) gamma ray spectrometer reduces the number of readings that can be made each day. In a mobile survey, increasing the counting time means that each measurement is taken over a longer distance and therefore each is representative of a larger sample volume. This smoothing effect may not be acceptable, and the vehicle speed may have to be reduced by a direct ratio to the increase in sample time. In the case of an airborne survey, this could reduce the aircraft speed to the point where the selection of a rotary wing aircraft is required, increasing the cost.

Alternative (2), increasing the size of an airborne detector package, may require a larger and consequently more costly type of aircraft. For portable or car-borne surveys, this may be a more economical solution. Alternative (3) moving closer to the source may be impossible, as in the case of a car-borne survey, or it may involve flying at unsafe altitudes for airborne surveys. Under alternative (3), the sample volume will also decrease appreciably and it may be necessary to make more closely spaced measurements for the same ground coverage. This also represents an increase in costs.

11.14.3.3 Instrument Parameters in GRS

Besides the myriad of other factors of concern in GRS, the instrumentation itself has several variables, which should be optimized for proper, accurate, and efficient performance. Some of the most important parameters and their meaning are described later.

11.14.3.3.1 Detectors used in Earth sciences

The most widely used scintillation detector material in Earth sciences is thallium-activated sodium iodide (NaI (Tl)) as mentioned earlier in Section 11.14.2.7.2 (scintillometers and scintillation detectors). For radiometric techniques used in geophysics, two other scintillation materials are also used but less often than NaI. These are cesium iodide activated with either thallium or sodium (CsI(Tl), CsI(Na)), and bismuth germanate ($Bi_4Ge_3O_{12}$), also called BGO. The CsI density is 4.51 g cm^{-3} making it about 25% more efficient at stopping gamma rays than NaI. Even more efficient is BGO with a density of 7.13 g cm^{-3}. However, the energy resolution of BGO is usually lower than NaI for use in spectrometry, but it is used in some handheld portable spectrometers where small size is important. The high efficiency also makes it attractive for use in gamma ray logging where the detector size is restricted by the borehole diameter. A relatively new scintillation material showing promise for use in radiometric techniques for

geophysics is lanthanum bromide ($LaBr_3(Ce)$) with density 5.08 g cm^{-3} and a resolution of <3%. However, at present, they can only be manufactured in relatively small sizes (e.g., 76 mm × 76 mm) and are extremely expensive.

Very high-resolution solid-state detectors (such as hyperpure germanium) make it possible to separate and measure gamma ray peaks of only slightly different energies, which are indistinguishable using sodium iodide or other scintillators. Daughter isotopes higher in the decay series and therefore less likely to be in disequilibrium can be used (Tanner et al., 1977). However, solid-state detectors are very small in size, which makes them less effective than scintillation detectors for spectrometric measurements of natural radiation in the field and restricts them for use in a laboratory environment. Modern airborne spectrometer systems are typically composed of a detector with at least two 16.4 l detector packages. Each detector package consists of four 10.2 cm × 10.2 cm × 40.6 cm (often quoted in inches: 4″ × 4″ × 16") thallium-doped NaI crystals in a thermally insulated container.

11.14.3.3.1.1 Properties of detectors

It is important to understand the properties of gamma ray detectors (scintillation detectors or crystals) to know which are the best and most appropriate for gamma ray spectrometric surveying. Some of the most important properties are the following:

- Light yield, measured in photons/keV, is the number of photons produced by the scintillator per keV of gamma ray energy. A high light yield is a good feature.
- Light output is given as a percentage of the light yield of NaI. When there is a low light output, it is difficult for the PMT to convert light to an electric pulse. The change of light output with temperature should be small or temperature changes will change the gain of the system and the measured gamma ray energies.
- Energy resolution was defined earlier in Section 11.14.2.7.2.1 as the ratio of the FWHM of the 662 keV gamma ray peak from Cs-137 to the energy of the peak. Large detectors for AGRS have energy resolutions ranging from 7% to 10% where lower values produce sharper peaks. For example, one manufacturer's instrument specifications quote a detector resolution of <8.5%. The energy resolution for any given detector is affected by the detector size. Examples from one manufacturer for 76 mm × 76 mm (3″ × 3") detectors are BGO = 9.7% and LaBr = 2.9% (lanthanum bromide).
- Density in g cm^{-3} is related to the ability to stop gamma rays. High-density scintillators stop (absorb) gamma rays in a shorter distance than low-density scintillators that is why BGO (the highest-density detector) is often chosen for use in restricted spaces like in borehole logging.

Some scintillation crystals are hygroscopic and must be hermetically sealed, or they will rapidly deteriorate due to absorption of moisture from the air. Table 4 gives some of the basic properties of scintillation detectors.

Some other exotic (today) scintillation materials with potential for future use in radiometric techniques for Earth sciences include BaF_2, barium fluoride; $CdWO_4$, cadmium tungstate; CaF_2 (Eu), europium-doped calcium fluoride; and plastic scintillators.

Table 4 Properties of scintillation detectors

Scintillator	*Light yield (photons keV^{-1})*	*Light output (% of NaI(Tl))*	*Density g (cm^3)$^{-1}$*
NaI(Tl)	38	100	3.67
CsI(Na)	41	85	4.51
BGO	8–10	20	7.13
$LaBr_3(Ce)$	63	165	5.08

11.14.3.3.2 Stabilization

In GRS, one of the first requirements is to determine the relationship between gamma ray energy and pulse height with a procedure called the energy calibration as described earlier (in Section 11.14.2.7.3.1). Once the energy calibration is completed, and the gain in keV/channel is set, the spectrometer is ready to record gamma ray counts. However, as mentioned earlier, due to the effect of temperature variations on the electronic components, the instrument may drift out of calibration and, in older systems, had to be rechecked periodically by the operator. One advance made in some instruments was to have a small radioactive source attached to the detector and its gamma ray peak could be continuously monitored by the spectrometer to keep it stabilized and properly calibrated. The source is called a stabilization source and was usually ^{133}Ba, which emits a gamma ray of about 350 keV.

An alternative to the use of a stabilization source is to control (stabilize) the temperature of the detectors and associated electronics. This was often done with large airborne detectors by packing them in an insulated box with a thermostatically controlled heater. The temperature of the box is maintained constant at some temperature that is sufficiently high that it will not be exceeded by ambient temperatures during the survey. Most modern systems still use an insulated box but without a heater. They continuously monitor the channel position of a prominent photopeak in the natural gamma ray spectrum from K, U, and Th and automatically adjust the gain to individual crystal PMTs of multidetector systems to maintain spectral stability. Even better are systems that monitor more than one peak for stabilization.

11.14.3.3.3 Pulse pileup

Besides dead time discussed earlier, an additional consideration at high count rates is pulse pileup and sum peaks or coincidence peaks. When two gamma rays strike the detector at the same time, or almost the same time, the two scintillations produced are 'seen' as one by the PMT, which produces a single pulse. Then, only one gamma ray count is registered, at an energy approximately equal to the sum of the energies of the two gamma rays, which struck the detector. For example, the 0.58 and 2.62 MeV gamma rays from the thorium decay series are both of high intensity and the sum peak at 3.20 MeV becomes significant for measurements of high thorium concentrations.

Pulse pileup also occurs when pulses from a detector arrive closer in time at the ADC than the pulse resolution time of the

ADC. When a pulse rides on the tail of the previous pulse, the system cannot measure the pulse heights correctly. If the pulses are very close in time, the system will record the two pulses as a single pulse with gamma ray energy of the sum of the pulse amplitudes called peak pileup. If the pulses are spaced further apart, the system may record them with incorrect pulse amplitudes, called tail pileup. In either case, the gamma rays will be counted in the wrong energy channels.

Pileup is only a problem at higher count rates when gamma rays are recorded in the wrong channels higher up in the spectrum leading to incorrect calculated radioelement concentrations. The number of pileup events depends on the count rate capability of the system. A good spectrometer has built-in pileup rejection to minimize the effect of pileup. The number of pileup events is smaller for ADCs with shorter pulse resolution time.

11.14.3.3.4 Spectrum linearity

Linearity was introduced in the discussion of energy calibration in the preceding text (Section 11.14.2.7.3.1). It refers to the relation between gamma ray energy (keV) and pulse height, or channel number in a multichannel spectrometer. The relation should be linear but several things can cause nonlinearity.

First, the light (scintillation) produced by a detector when it is struck by a gamma ray is converted to an electronic pulse by the PMT, which is then measured by the ADC to count it in the appropriate channel of the spectrum. However, the light, and therefore the electronic pulse created from scintillations in a detector, is not a perfect linear function of the gamma ray energy deposited in the crystal, resulting in a nonlinear relationship between channels and energy. Secondly, electronics can also affect the linearity but this should be minimal if they are of good quality. The nonlinearity of crystal output can be of the order of 2–4%. In large airborne systems, the output of several detectors is often summed for input to the ADC. However, different crystals can have different nonlinear response curves, and after summing the output of different detectors, the sum spectrum can be distorted resulting in poor energy resolution.

Modern spectrometer systems can correct for the nonlinear effects in the crystals in multidetector systems. Ideally, each crystal should be corrected independently to ensure that individual crystal nonlinearity is properly corrected. The different methods to correct for the crystal nonlinearity are typically proprietary. If the crystal linearization is carried out correctly, the output spectrum will have a close to perfect linear relationship between channels and energy, that is, the conversion between channels and energy will be linear (e.g., 3 keV/channel or 6 keV/channel).

11.14.3.3.5 Dead time

The equipment dead time (t) of a gamma ray spectrometer is effectively the time taken by the equipment to analyze a single gamma ray that has been detected. During this period, the equipment is busy and cannot analyze any other gamma rays. This dead time (sometimes called resolving time) is only a few microseconds per recorded count, but it becomes an appreciable percentage of the counting time at high count rates and must be taken into consideration. In a given sample time of say 1 s, the dead time is the equipment dead time multiplied by the total number of pulses C recorded in that 1 s. Thus, during the 1 s sample time, the time available for the spectrometer to count gamma rays is the 'live time.' The dead time is then the sample time minus the live time.

The true count N can be computed from the observed count (or count rate) n if the equipment dead time t is known, using the following equation (IAEA, 2003):

$$N = n(1 - Ct)^{-1} \quad [9a]$$

or alternatively,

$$N = n\ (\text{live time})^{-1} \quad [9b]$$

where N is the corrected count in each second (i.e., count rate), n is the raw count recorded in each second, C is the TC over all channels in 1 s, and t is the equipment dead time per pulse.

The product $C \times t$ is the recorded dead time, the time taken to process all pulses reaching the detector over all channels in 1 s (i.e., total count × equipment dead time per pulse).

Dead time should be kept at a minimum, preferably below 10%, by suitably adjusting detector sizes, source–detector spacing, or other variables. Methods of computing or measuring the equipment dead time have been given by IAEA (2003) and Chase and Rabinowitz (1968) and, for a borehole gamma ray logging system, by Crew and Berkoff (1970).

Different manufacturers define dead time slightly differently. One method is to use two 'clocks,' one that runs continuously measuring real time and a 'live time' clock that is stopped whenever the system is busy processing a pulse. The difference between the clock times is the 'dead time,' or the time when the spectrometer was busy. Often, dead time is defined as the effective count loss expressed as a percent. After a dead time correction has been applied, the corrected count rate should be the count rate for the equivalent dead time-free system. The latest most advanced spectrometers have virtually zero dead time.

11.14.3.4 Some Features of New Systems

The newest airborne gamma ray spectrometric systems feature automatic gain stabilization, more compact packaging, and energy calibration using the natural radioelements K, U, and Th and an Ethernet connection between components, which also permits 'remote access' by engineers (via telephone modem) to analyze systems and check their performance from headquarters.

Other major new features include an individual ADC for each detector crystal and an energy linearization system that converts incoming nonlinear data from the detector to a fully linear spectrum. This allows spectrometers to process signals from each detector separately instead of summing the signals from all detectors. Software has also been developed to take advantage of the full spectral recording instead of using only a few windows in the spectrum.

About 1024 channels are now being recorded routinely by many survey contractors. Sampling rates have increased, often ten times per second, which means recording an entire 1024 channel spectrum every 0.1 s. Larger detector volumes of up to 80 l are being flown for higher sensitivity and upward-looking

background detectors (see Section 11.14.5.3.6.1) are now common. The use of cesium iodide (CsI) for detectors instead of sodium iodide (NaI) is being tested as a possible improvement.

11.14.3.5 Calibration for Quantitative Radioelement Measurements

In the early days of GRS, the results were reported in cps in the various windows. This was not satisfactory for quantitative measurements. To make meaningful radioelement measurements with a gamma ray spectrometer, it must be calibrated with respect to some standard. Without calibration, confusion can result when comparing gamma ray counting measurements made with different instruments, sometimes even when the instruments are of the same manufacture. In addition, different instruments may display the data in different units, such as cps, microroentgens per hour or millivolts. The International Atomic Energy Agency (IAEA, 1976) formulated recommendations for standardizing measurements in exploration by calibration and presenting results in units of concentration (e.g., %K, ppm eU, and ppm eTh).

The measured radioactivity (gamma ray intensity) is proportional to the radioelement concentration in the source of the radiation. The gamma ray intensity must first be corrected for background and for interference from gamma rays from other radioelements than the one being measured. Then,

$$\text{Radioelement content} = (\text{a constant}) \times (\text{gamma ray intensity}) \qquad [10]$$

The major component of calibration relates to the determination of this constant of proportionality, which is the '*S* factor' or the reciprocal of the sensitivity (also sometimes called K, or *C* factor). To determine the *S* factor or the sensitivity of each individual radioelement K, U, and Th, the gamma ray counts due to each individual radioelement must be determined. This is accomplished by the procedure known as spectral stripping (see Sections 11.14.2.8.1, 11.14.4.2.4.1, and 11.14.5.3.8), which is applicable in surface, airborne, and borehole measurements. The sensitivity has units of count rate per unit of radioelement concentration. The rigorous solution for the calibration equations is given by Grasty (1979) and in IAEA (1991, 2003).

To summarize, calibration for in situ radioelement assays refers to the determination of the calibration factors (stripping ratios and sensitivities) necessary to convert the count rates measured over soil and rocks with airborne, surface, or borehole gamma ray spectrometers into concentrations of K, U, and Th in the ground (usually % K and ppm U and ppm Th). These calibration constants are specific to the instrument and to the sample geometry for which it was calibrated (usually 2π geometry, which is a 'flat' surface or 4π in borehole logging). The calibration is done, in part, on concrete calibration pads for airborne and surface spectrometers such as the set of pads constructed by the GSC, located at the Ottawa Airport (Darnley, 1970). Other sets of calibration pads (IAEA, 1989) are located at airports around the world, such as in Thailand and in the United States, as well as at other nonairport locations of calibration facilities for portable or borehole spectrometers (concrete pads or concrete model boreholes).

Finally, for airborne spectrometers, the rate of change of these parameters with height must be determined to be able to apply corrections to the recorded airborne data as the height of the aircraft above ground changes. Usually, all the calibration factors are precomputed for flying at say 120 m above ground, and then, corrections are applied for the small deviations above and below this survey height. Other corrections also to be determined as part of calibration are required, such as corrections for casing and fluid in the hole for borehole spectrometers and cosmic ray corrections for airborne spectrometers. Calibration facilities are described in the succeeding text.

11.14.3.6 Calibration Facilities

11.14.3.6.1 Introduction

Calibration or standardization of gamma ray spectrometer equipment is based on a known, primary reference standard. Secondary calibration standards are then derived from the primary reference. Ultimately, all measurements depend on calibration of a laboratory gamma ray spectrometer using standard samples. The laboratory spectrometer analyzes samples taken during construction of concrete calibration pads, and concrete model holes, and samples taken from test 'strips' or 'ranges.' The concrete models become secondary calibration sources for field gamma ray spectrometers. The calibration of an AGRS requires, in addition to the concrete calibration pads, a calibration test 'strip' or 'range' over which test flights may be made. The K, U, and Th analysis of this test strip is done by in situ assays on a grid with a calibrated portable gamma ray spectrometer. Once an instrument is calibrated, it should be recalibrated periodically because its characteristics may change with time.

The following are recommendations for a test strip (IAEA, 1976):

1. It should be flat (low topographic relief).
2. It should have uniform radioactivity (i.e., no anomalous values, as tested by surface traverses and sampling).
3. It should have minimum dimensions of 1.0 km × 3.0 km.
4. It should be close to a large body of water for background measurements.
5. It should have easily repeatable navigability.

Calibration facilities (model boreholes, also called 'test pits') for TC gamma ray logging were first made available in the United States in the early 1960s in Grand Junction, Colorado, with an additional model hole each for K, U, and Th for gamma ray spectral logging being built in 1974 (Mathews et al., 1978). The first five large calibration pads for calibrating airborne and ground gamma ray spectrometer (GGRS) systems were built in Ottawa in 1968 (Darnley, 1970; Grasty and Darnley, 1971), and five large pads were built at the airport in Grand Junction, Colorado, in 1976 (NURE, 1976). Four calibration pads built specifically for portable gamma ray spectrometers were constructed at Risø in Denmark in 1971 (Lovborg and Kirkegaard, 1974; Lovborg et al., 1972, 1978a,b). Radiometric test ranges (test strips) were established at Breckenridge near Ottawa in 1970 (Charbonneau and Darnley, 1970; Grasty and

Charbonneau, 1974) and in the United States near Lake Mead, Arizona (NURE, 1976). Extensive calibration facilities for borehole gamma ray spectral logging (3 radioelements, 9 models, and 27 holes) and for surface portable GRS (10 pads) were constructed in Ottawa in 1977 (Killeen, 1978, 1986a,b; Killeen and Conaway, 1978). Calibration pads and model boreholes were also constructed in Canada in Calgary, Alberta, and Saskatoon, Saskatchewan, respectively (Killeen, 1986a).

Tables of radioelement (K, U, and Th) concentrations for all known (to that date) worldwide calibration facilities were given by Killeen (1979). These included calibration pads and test ranges (strips) for airborne (Canada, the United States, and Iran) and portable gamma ray spectrometers (Canada, Denmark, South Africa (Toens et al., 1973), and the United States) and model boreholes for borehole gamma ray spectrometers (Canada, South Africa, and the United States). Since these early 'heady' days of calibration facility construction, additional calibration pads and model holes have been constructed in many parts of the world as nations implemented the recommendations of the IAEA (1976) (e.g., Denmark, Czech Republic, Hungary, Greece, Argentina, Egypt, India, Thailand, Brazil, Spain, and Australia (Killeen and Elliott, 1990)). Pads for calibration of AGRS were constructed in Sweden using assemblies of precast concrete blocks fitted together to form large pads. Construction of model boreholes was also completed in Sweden (Osterlund, 1986), and a set of four pads, each 8 m square, was constructed at an airport near Helsinki, Finland (Killeen, 1979). The design, construction, and use of a set of 7.6 m square pads in Thailand for AGRS were given by Grasty (1987). Model boreholes for gamma ray logging were also constructed in Houston by the American Petroleum Institute (API) for use in oil well logging (Belknap et al., 1959). More recently, smaller transportable model boreholes have been developed (Grasty and Blanchard, 1993). A map showing the worldwide distribution of gamma ray spectrometer calibration facilities is given in IAEA (2010), and a directory of all known gamma ray spectrometric calibration sites and the address of the lead agencies associated with them are also given in appendices.

11.14.3.6.2 Calibration pads and model boreholes

The calibration of an AGRS system requires calibration facilities with both concrete calibration pads and an airborne calibration test strip. The aircraft can be towed or taxied under its own power to a set of concrete calibration pads such as those built by the Geological Survey of Canada at Ottawa or those built by the US Department of Energy at Grand Junction, Colorado (or to smaller transportable pads with correction for percentage of infinite source).

Pad dimensions in the early calibration facility construction were nominally 8 m × 8 m or larger to simulate an infinite 2π surface (as much as is feasible) and about 50 cm thick (i.e., infinitely thick) and were spaced 15 m apart or greater. The concentrations of the radioelements in the concrete mixtures of the calibration pads are controlled by adding natural minerals containing K, U, or Th to produce a K pad, U pad, and Th pad plus a 'blank' pad, which uses the same concrete but without any K, U, or Th added to the mix. Sometimes, a fifth pad was constructed for surface and airborne calibration, with a mixture of the K, U, and Th additives. The radioelement concentrations of the pads are relatively low because they are designed for calibrating AGRS with large detector volumes. Typical maximum concentrations are about 2–5% K, 10–30 ppm eU, and 25–45 ppm eTh.

For portable (surface) gamma ray spectrometers, calibration pads were from 50 to 60 cm thick with dimensions from 1 to 3 m diameter and were cylindrical or square (IAEA, 1989). A minimum of four pads are required (K, U, Th, and blank). The radioelement concentrations of the pads can be higher than for airborne calibration pads because portable gamma ray spectrometers have relatively smaller detector volumes. Typical maximum concentration ranges are from 2% to 7% K, 350 to 3000 ppm eU, and 300 to 4000 ppm eTh.

Although these relatively large pads for airborne and surface spectrometers are still in use today, they have generally been replaced by smaller pads of dimensions 1 m × 1 m × 0.3 m, which are 'transportable' and usually mounted on a wooden pallet. For calibration, the pads are mounted on a metal trolley, which is rolled into a reproducible position (for each of the four pad measurements) below the detectors, which are inside an aircraft (Grasty et al., 1991). The same pads can be used for calibrating portable spectrometers. In this case, a geometric correction (described in **Section 11.14.4.2.4**) is applied to the derived sensitivities to account for the percent of an infinite source represented by the pad (IAEA, 2003).

For calibration of borehole gamma ray spectrometers, three model boreholes are constructed consisting of a vertical concrete column or cylinder of 1 m diameter or more, with a simulated ore zone from 1.0 to 1.5 m thick sandwiched between upper and lower barren zones. Each column contains a borehole of diameter the same as expected in the field (e.g., 75 mm) intersecting the ore zone. Three columns are required, which contain 'ore' zones for potassium, for uranium, and for thorium. The barren zones serve as the blank zone for the calibration. Often, more than one borehole model is constructed with different concentrations for each of K, U, and Th (Killeen and Conaway, 1978). The range of radioelement concentrations in the ore zones is usually similar to that for pads used to calibrate portable gamma ray spectrometers. With the more recent discovery of very high grade uranium deposits, models with much higher U grades have been constructed such as those in Saskatoon, Saskatchewan, Canada (Killeen, 1986a).

Besides the previously mentioned man-made calibration facilities, radiometric test ranges were made available for AGRS. In Canada, the Breckenridge test strip, which was established by the GSC, is located on the Ottawa River floodplain near Ottawa. It is 7 km long, close to the 1 km wide river, with a railroad track along the strip, which is a convenient aid to navigation and facilitates reproducible results (Charbonneau and Darnley, 1970; Grasty and Charbonneau, 1974). In the United States, a radiometric test range referred to as the 'Lake Mead Dynamic Test Range' is located about 75 km due east of Las Vegas, Nevada (NURE, 1976).

11.14.3.6.3 Construction of calibration facilities

The main points regarding the construction of the earlier calibration facilities were summarized in a very instructive report (IAEA, 1976). Some data on the details of the actual construction procedures for the calibration pads in Denmark have been described (Lovborg et al., 1972, 1978a,b). The construction of

the large calibration pads for airborne systems was described by Grasty and Darnley (1971) for the pads at Ottawa and by Ward (1978) for the pads at Grand Junction. More recent construction information can be found in Grasty (1987), Grasty and Blanchard (1993), Grasty et al. (1991), and IAEA (1989).

It is important to understand and appreciate the relationship between primary and secondary calibration facilities for various types of gamma ray spectrometers. First, international standard samples are used to calibrate laboratory spectrometers. These in turn analyze samples from calibration pads and model boreholes during their construction, which can then be used as secondary calibration facilities for field spectrometers. The calibrated field spectrometers can then carry out radioelement assays in situ.

Therefore, all calibration relies on some internationally agreed upon 'standards.' These are usually specially selected samples prepared in bulk, analyzed by several methods by several laboratories, and distributed by some recognized laboratory, which maintains a large quantity of the standard for distribution, for a fee. For example, the preparation and certification of reference materials for GRS are described in IAEA (1987). It includes standards for U, Th, and K determinations. These reference materials came from Canada and were originally prepared in Ottawa at Natural Resources Canada, Canadian Certified Reference Materials Project, Mineral Sciences Laboratories, Canada Centre for Mining and Energy Technology (CANMET). In the United States, samples used for standards are available in Argonne, Illinois, from the Standards and Reference Materials section, US Department of Energy, New Brunswick Laboratory.

It is also important that the samples be in radioactive equilibrium or that the state of disequilibrium be known since the samples will be used to calibrate a radiometric method of analysis rather than a chemical method. When international standards are not used, problems can arise when spectrometers are calibrated in different calibration facilities in different countries. An example of this was documented in an international intercalibration project for borehole gamma ray logging sponsored by the IAEA (Killeen and Elliott, 1990).

11.14.4 Data Acquisition

11.14.4.1 Introduction

Radiometric surveys on the ground surface are primarily 'passive,' measuring natural radioactivity. However, some 'active' methods are in use on the ground surface and in boreholes. These are more appropriately called 'nuclear techniques' (IAEA, 1977; Morse, 1977). The term 'radiometric techniques' is generally considered to refer to measurements of natural radiation. One example of an active method is elemental analysis with portable x-ray fluorescence (XRF) instruments. Portable XRF instruments use a shielded radioactive source to bombard the sample (e.g., rock outcrop) with a collimated beam of gamma rays when a shutter is opened, by pressing the detector against the rock surface. The gamma rays excite the elements in the rock, which then emit x-rays with energies characteristic of the emitting element. The detected x-rays are analyzed by the instrument in a similar process to GRS. Early car-borne experiments were carried out with a neutron activation analysis system mounted on a trailer (Senftle and Hoyte, 1966), but this proved to be impractical primarily due to radiation safety concerns. However, active methods are used frequently in borehole logging, primarily in hydrocarbon exploration and to a lesser extent in mineral exploration. These mainly include the use of gamma ray sources or neutron sources in a borehole probe, which bombard the wall of the hole to derive information about the density (gamma–gamma density probe) or porosity (neutron porosity probe) of the wall rocks. The use of nuclear sources in borehole logging for mineral exploration applications was described by Killeen (1997b), including density logging and spectral gamma–gamma logging for heavy element detection, and the location of ore zones and determination of their grade. A version of a system for borehole elemental assaying using neutron activation analysis was also developed for use in mining (Nargolwalla, 1975). The passive methods are by far the most widely used radiometric techniques, in particular the use of gamma ray scintillometers and gamma ray spectrometers, which are the main subject of this chapter. This data acquisition section is primarily about those passive methods.

The application of GRS in the field using surface, airborne, and borehole equipment is based on the same fundamental principles but differs slightly in the derivation of calibration factors and correction factors and their application during field measurements (data acquisition) and data processing. The following describes the methodology that is common among the three types of measurements as well as the differences that must be taken into consideration during both calibration and field surveys. Surface portable and car-borne gamma ray spectrometers are considered first, followed by AGRS, which have some of the calibration procedures in common with portable spectrometers, and finally, borehole gamma ray and gamma ray spectral logging are covered.

11.14.4.2 Data Acquisition with Surface Systems

11.14.4.2.1 Portable gamma ray spectrometers

For field measurements, with a surface portable gamma ray spectrometer, the detector may be carried in a backpack, but it is usually placed on the ground surface during the measurement. The height of the detector aboveground affects the size of the sample volume. In any case, the measurement method should be consistent so that changes in geometry do not change the results. Depending on the objective of the survey, measurements may be made at a regular spacing along a survey line, or on a regular grid, or as spot measurements on sources of interest such as in situ assays on rock outcrops. Often in this case, a scintillometer is first used to quickly scan the outcrop for homogeneity, to locate an area of average radioactivity for the assay, ensuring it is not a measurement on an anomaly. A typical portable spectrometer would have a 76 mm × 76 mm NaI (Tl) detector with a volume of 350 cm^3. Some smaller spectrometers use CsI or BGO detectors, which are more efficient than NaI. It may be a four-channel (window) spectrometer or a multichannel (e.g., 256) spectrometer from which the four window counts are computed and displayed. Counting times for measurements are varied according to the count rate, to obtain good counting statistics. In areas of low radioactivity, longer counting times are used than in areas of high radioactivity.

11.14.4.2.2 Car-borne gamma ray spectrometers

Field measurements with a car-borne gamma ray spectrometer can be made with the detector inside the vehicle or mounted on the roof if a larger sample volume is desired. In a recent environmental survey, large-volume NaI(Tl) detector packs (as used in airborne surveys) were mounted on a platform carried on the front lift of a small tractor (Killeen, 2013). Again, depending on the objective of the survey, measurements may be made at regular intervals along roads or on a regular grid of roads or off-road to avoid having the roadbed materials dominate the measurement. In some environmental surveys such as in a 'radioactive spill,' the road itself may be the subject of investigation. Usually, the detector is quite large, often of the type and size used in airborne spectrometers. Manufacturers have developed multichannel systems with the flexibility to be easily installed in a helicopter or in a ground vehicle and the data acquisition (recording) is much the same.

11.14.4.2.3 Background radiation in surface measurements

There are several ways in which the background can be determined for both car-borne and portable gamma ray spectrometers. Ideally, a measurement can be made over water in the survey area using a boat. This is relatively easy with a portable spectrometer. For a car-borne spectrometer, which typically has a larger detector, it is suggested that the detector be connected to a portable spectrometer to measure background from a boat (Grasty and Cox, 1997). Another method is by driving the vehicle onto a bridge (not stone or concrete) over a wide river or onto the ice of a lake (Bowie et al., 1955). This will yield a background value for the equipment itself, the vehicle, cosmic rays, and radioactivity in the atmosphere.

11.14.4.2.4 Calibration of portable gamma ray spectrometers

It is important to simulate the field geometry as much as possible during the calibration of a spectrometer. For a portable gamma ray spectrometer, the field measurements are usually made in a 2π geometry (see **Figure 7**) by placing the detectors directly on the ground surface, and for the calibration measurement, the detectors are placed directly on the calibration pads. The sample time or counting time may be as long as required to provide good counting statistics. The complete calibration can be carried out by making measurements as described in the succeeding text, on a set of four calibration pads such as those recommended by the IAEA (1976, 1989, 1991, 2003). Early pads were circular and 3 m in diameter, but in recent years, smaller pads are being used with geometric corrections to compensate for the pads not being infinite.

Even if the detector is to be located at a raised elevation with respect to the ground, such as when mounted on a backpack or a vehicle, the system should first be calibrated with the detector on or near the surface of the calibration pad to ensure that the pad represents as near as possible an infinite 2π geometry source. Otherwise, appropriate geometric corrections must be applied to the derived sensitivities to account for the smaller percentage of an infinite source that the pad represents as the detector height h above the pad increases. The ratio R of gamma radiation from a cylindrical pad 0.5 m thick to that from an infinite source with the detector placed on the center of the pad is (IAEA, 2003)

$$R = 1 - hr^{-1} \quad \left(\text{if } hr^{-1} < 0.2\right)\infty \qquad [11]$$

where h is the height of the center of the scintillation detector crystal above the pad in meters and r is the radius of the pad in meters.

The geometric correction G, which is R^{-1}, is used to multiply the sensitivities (which are K, U, and Th window count rates, per unit of radioelement concentration (% K, ppm U, and ppm Th)), derived from the calibration procedure, to obtain the corrected sensitivities. In the field, measurements with the detector mounted at different heights will be representative of a larger sample volume (IAEA, 1979, 1989).

The center of the detector for most portable spectrometers is about 6 cm above the pad. Actual geometric corrections depend on several other factors including pad density, thickness, and the window energy. For a detector 6 cm above a 0.5 m thick circular pad, corrections are typically about 1.045 for a 3 m pad and 1.07 for a 2 m pad. For a 0.3 m thick, 1.0 m^2 'transportable' pad, the correction is about 1.17 (Grasty et al., 1991; IAEA, 2003). An accurate method to compute geometric corrections taking the other factors into consideration was given by Lovborg et al. (1972).

11.14.4.2.4.1 Determination of stripping ratios and sensitivities

Calibration of portable gamma ray spectrometers is carried out (as recommended by the IAEA, 1991) by a series of measurements on concrete calibration pads of thickness 0.3–0.5 m and diameters from 1.0 to 3.0 m, designed to simulate rocks (outcrops) expected in the field measurements. For portable spectrometers, measurements on the calibration pads can provide a complete calibration with the determination of both the stripping ratios and the sensitivities. The same measurements also form a part of the calibration of car-borne and airborne systems, but for these systems, only the stripping ratios can be determined with the pads, not the sensitivities. Determination of these sensitivities will be described in the next section. Ideally, the pads are pure sources of K, U, and Th produced by adding to the concrete mix potassium, uranium, and thorium minerals to provide three 'monoelemental' sources, one each for K, U, and Th and in addition, a fourth pad called the 'blank,' which is the same concrete without any additives. The blank pad is used for measurement of the background radiation, which can then be subtracted from the measurements on the K, U, and Th pads. The spectrometer measurements can be used to derive the stripping ratios and sensitivities as described in IAEA (1976, 1991, 2003), Lovborg (1984), and Grasty and Minty (1995) and summarized later.

The spectrometer measurements provide K, U, and Th window counts (or count rates) on each of the four pads. The four pads have documented K, U, and Th concentrations based on samples taken during construction, which were analyzed with a laboratory spectrometer. The count rates n_K, n_U, and n_{Th} on any pad are linearly related to the K, U, and Th concentrations of the pad, C_K, C_U, and C_{Th}, and the backgrounds b_K, b_U, and b_{Th} with the following equations (IAEA, 1991):

$$n_K = S_{K,K}C_K + S_{K,U}C_U + S_{K,Th}C_{Th} + b_K \qquad [12]$$

$$n_U = S_{U,K}C_K + S_{U,U}C_U + S_{U,Th}C_U + b_U \qquad [13]$$

$$n_{Th} = S_{Th,K}C_K + S_{Th,U}C_U + S_{Th,Th}C_{Th} + b_{Th} \qquad [14]$$

These can be simplified to three equations with three unknowns by subtracting the background window counts from

the window counts n_K, n_U, and n_{Th} and adjusting the pad concentrations by subtracting the blank (background) pad concentrations from the K, U, and Th concentrations C_K, C_U, and C_{Th}. This becomes a 3×3 matrix equation (Lovborg et al., 1972)

$$\mathbf{N} = \mathbf{SC} \quad [15]$$

where **N** is a matrix of background-corrected count rates on the K, U, and Th pads; **S** is a matrix of sensitivities; and **C** is a matrix of the adjusted C_K, C_U, and C_{Th} concentrations.

The sensitivity matrix is determined using

$$\mathbf{S} = \mathbf{NC}^{-1} \quad [16]$$

where $\mathbf{C}^{-1}$ is the inverse of **C**.

The radioelement concentration (or 'assay') sensitivities are given by $S_{K,K}$, $S_{U,U}$, and $S_{Th,Th}$, which are measured as counts in the K, U, and Th windows per unit concentration of K, U, and Th, respectively. The stripping ratios, which were described earlier, are given by

$$\alpha = \frac{S_{U,Th}}{S_{Th,Th}} \quad [17]$$

$$\beta = \frac{S_{K,Th}}{S_{Th,Th}} \quad [18]$$

$$\gamma = \frac{S_{K,U}}{S_{U,U}} \quad [19]$$

$$a = \frac{S_{Th,U}}{S_{U,U}} \quad [20]$$

$$b = \frac{S_{Th,K}}{S_{K,K}} \quad [21]$$

$$g = \frac{S_{U,K}}{S_{K,K}} \quad [22]$$

The stripping ratios can now be used to strip the counts in the K, U, and Th windows using the stripping eqns [30]–[33] given in the data processing Section 11.14.5.2.1. (Note: some modern portable gamma ray spectrometers are sold 'precalibrated.')

11.14.4.2.5 Determination of sensitivities for car-borne and airborne systems

For portable spectrometers, when the count rates recorded in the K, U, and Th windows have been stripped (and the backgrounds subtracted), these stripped counts can be used to determine the concentration of K, U, and Th on the ground in rocks and soils, with the S factors (sensitivities) for each window, all determined from calibration pad measurements described earlier (for 2π geometry and assuming radioactive equilibrium). However, car-borne spectrometers require a 15–20 m diameter calibration area (IAEA, 2003), and airborne spectrometers require a calibration test strip to determine the sensitivities. The calibration area and test strip have known radioelement concentrations measured with a calibrated portable gamma ray spectrometer, which are then related to the fully corrected car-borne or airborne spectrometer count rates using eqn [23] to determine the sensitivities:

$$S = NC^{-1} \quad [23]$$

where S is the sensitivity for the K, U, or Th window (in units of cps $\times$ (ppm or %) $^{-1}$); N is the count rate measured in the test area or strip for each K, U, or Th window after dead time, background, and stripping corrections; and C is the concentration of radioelements (K%, eU ppm, or eTh ppm) in the test area or strip.

More details on the data acquisition over test strips to determine sensitivities for airborne systems, which must also include air attenuation corrections, are given in Sections 11.14.4.3.4 and 11.14.4.3.6. The calculated potassium, uranium, and thorium concentrations can then be determined from field survey measurements as described in the data processing Section 11.14.5.2.1, by dividing the measured, fully corrected K, U, or Th window count rate (N) by the appropriate K, U, or Th sensitivity (S).

11.14.4.3 Data Acquisition with Airborne Systems

11.14.4.3.1 Introduction

Airborne radiometric techniques in use today are primarily AGRS surveys with very few TC scintillometer surveys, which are now being pretty much a thing of the past. Here, data acquisition for AGRS is covered in detail.

Descriptions of modern AGRS methods can be found in IAEA (1991), in IAEA (2003), or in the Australian Geological Survey Organization (AGSO) publication (Grasty and Minty, 1995).

11.14.4.3.2 Airborne gamma ray spectrometric surveying

Airborne geophysical surveys are normally flown in a direction perpendicular to the main geologic strike of the survey area. For AGRS surveying, the line spacing depends on the objectives of the survey: In reconnaissance surveys, it may vary between 2000 and 5000 m; in regional surveys, it may vary between 500 and 2000 m; and in detailed surveys flown for mineral exploration purposes, flight line spacing can be as little as 100 or 200 m. The aboveground flight elevation (terrain clearance) in fixed-wing spectrometric surveys is normally 120 m, although it may vary between 100 and 150 m. For helicopter-borne surveys, the detectors are mounted in the helicopter, and the nominal terrain clearance varies between 60 and 120 m, depending on local terrain conditions and the configuration of other survey equipment. Typically, aircraft fly at a speed of approximately 120 knots (190 km h^{-1}).

11.14.4.3.3 The airborne gamma ray spectrometer requirements

AGRS surveys may be flown with small detectors in a very light aircraft or with a very large detector package in a larger aircraft, but in any case, the requirements for data acquisition, calibration, correction factors, and processing, to obtain accurate results, are the same. We will start by looking at data acquisition and processing from the viewpoint of requirements for an airborne gamma ray spectrometric survey for a large organization such as the GSC or AGSO. Typical specifications would usually require a gamma ray spectrometer system capable of recording at least 256 channels (usually 1024) of gamma ray data from 0.0 to 3.0 MeV, thermally insulated detector packages to minimize drift, on-line monitoring of a selected photopeak or photopeaks to stabilize gain, and an upward-looking detector (see Section 11.14.5.3.6.1) to monitor atmospheric radon changes (or some alternate approved method). Detector

size is often specified as 50.4 l of sodium iodide (4π) for the main detector and 8.4 l for the upward-looking (2π) detector (shielded by the main detector), which records uranium window counts. Maximum dead time for either the main or upward-looking PHA must not exceed 8 μs per pulse. Overall signal conditioning must yield a detector resolution better than 12% FWHM based on the 662 keV peak of ^{137}Cs.

The recorded spectral data produced by the system must be suitable for processing using the noise-adjusted singular value decomposition (NASVD) method (see **Section 11.14.5.3.1**), which requires that the Poisson nature of the recorded channel counts be retained to ensure correct spectral component analysis.

The pulse height analysis system used must be sufficiently linear to ensure that for all peak energies between ^{137}Cs at 662 keV and ^{208}Tl at 2615 keV, the relationship of energy to channel number (e.g., 3.0 keV/channel), or gain, will remain acceptably constant. For older spectrometers (still in use today), this meant a keV/channel variation of about ±1%; however, the gain in the latest spectrometers is essentially constant over the entire energy range of interest.

The system must permit cosmic ray activity to be monitored including all energies above 3000 keV (called the cosmic window). Often, the cosmic count is assigned to the final spectrum channel, for example, channel 256 or 1024. Channels 1–255 or channels 1–1023 are for potassium, uranium, and thorium measurements.

The following spectrometer data must be recorded digitally for each 1 s counting interval: the full spectrum (e.g., 1024 channels) from the main detector including the cosmic ray channel, the system dead time (or live time) to an accuracy of 1 ms, and the gamma ray energy windows in **Table 5** with limits given in keV from IAEA recommendations (IAEA, 1991, 2003).

Five other parameters, which must also be recorded (minimum interval or sensitivities), include radar altimeter (5 m), barometric altimeter (5 m), time (0.1 s), navigation output (10 m), and temperature (1 °C). Strictly speaking, barometric pressure is required, which is calculated from the uncompensated barometric altitude.

This information is used to correct gamma ray data for the height of the detector above the ground and correct for the effect of temperature, pressure, and air density on the air attenuation.

11.14.4.3.4 Calibration of airborne gamma ray spectrometers

The following information follows the approved calibration procedures for an AGRS survey for the GSC. As such, it is representative of an AGRS for similar government organizations or a large mining or oil company. The survey would typically have quality control overseen by a 'technical authority' who participates in all phases of the presurvey (or postsurvey) season tests.

Table 5 Gamma ray spectrometry energy windows

Name	*Element*	*Peak (keV)*	*Energy range (keV)*
Potassium	^{40}K	1460	1360–1560
Uranium	^{214}Bi	1760	1660–1860
Thorium	^{208}Tl	2615	2410–2810
Total count			410–2810
Cosmic			3000–6000
U (Up detector)	^{214}Bi	1760	1660–1860

Source: IAEA (1991) Airborne gamma ray spectrometer surveying. *Technical Reports Series No. 323*, p. 97. Vienna: International Atomic Energy Agency.

Prior to the survey, the system must be calibrated on a set of calibration pads (built to IAEA recommendations) to determine stripping ratios as described earlier for portable gamma ray spectrometers in **Section 11.14.4.2.4.1**. Calibration will be carried out with the detectors installed in the survey aircraft. To determine sensitivities for the K, U, and Th windows, presurvey-season test flights over a calibration range such as the Breckenridge test strip must be carried out, as well as the over-water background line, with the approval of the technical authority who will make arrangements for a ground survey of the test strip to be completed using a calibrated portable gamma ray spectrometer on the same day as the test flights. The results from these test flights will be used to determine the attenuation coefficients and radioelement sensitivities for the system. The energy resolution and energy calibration will also be established using ^{137}Cs during the preseason testing and monitored daily during the survey. The radar altimeter and system barometer must also be calibrated to the Global Positioning System (GPS) altimeter at the beginning of the season. This may be accomplished by flying over water or a suitable airport runway of known elevation at 30 m intervals for 100–200 s each between 60 and 300 m above the surface.

The cosmic window is calibrated by flying at 500 m intervals for 600 s each between 1500 and 3500 m above sea level (ASL) during the preseason test strip flights.

In summary, calibration data can be acquired as recommended by the IAEA (2003) as follows:

1. For the aircraft and cosmic background corrections, the aircraft flies a high-altitude stack (series of lines) over water at a series of different heights. Cosmic calibration need not actually be performed over water, although it is preferable. The requirement is to fly at altitudes sufficiently high that terrestrial radiation is negligible.
2. For the Compton stripping ratios, data are gathered by placing the individual spectrometer detector packs over specially constructed calibration pads for several minutes. These data determine the stripping ratios for the spectrometer.
3. For the height attenuation and factors for conversion of counts to radioelement content (sensitivities), the aircraft flies an altitude stack over a nominated line over a dynamic test area, at a series of heights, which would be encountered during a typical radiometric survey.
4. For the radon background correction, the aircraft collects data over water with little or no radon, and again when there is significant radon over water (based on multiple flights over water).

11.14.4.3.5 Determination of attenuation coefficients

Gamma radiation is attenuated by air, and over an infinite source, this attenuation can be approximated by an exponential function. Determination of the infinite source attenuation coefficients is accomplished by flying over the Breckenridge test strip (or other calibration range) and over water

(to measure the overwater background) at 30 m altitude intervals for 100–200 s each, typically between 60 and 300 m above the surface. For each flight height over the test strip and over water, the mean background count rates in the TC, K, U, and Th windows are subtracted from the mean window count rates over the test strip. The effective STP (standard temperature and pressure) height is then determined from the average air temperature and pressure at each flight height as described in Section 11.14.5.3.7. Next, the stripping ratios are computed for each effective STP height, followed by computation of the stripped window counts at each effective STP height. Finally, the height attenuation coefficients for each of the four windows are determined by fitting the stripped count rates at each altitude to an exponential function for each of the four windows.

11.14.4.3.6 Determination of window sensitivities

The window sensitivities (counts per unit concentration of K, U, and Th) for the planned (nominal) survey altitude are determined from the data recorded in the flights at different altitudes over the calibration test strip and over water, which were used to determine the attenuation coefficients. The ground concentrations of K, U, and Th on the calibration test strip must be measured with a calibrated portable gamma ray spectrometer on the same day (or close to it) as the test flights to ensure that the soil moisture is the same for both the airborne and ground measurements. The sensitivity for each window is determined from the relation between the stripped window counts at planned survey altitude and the measured ground concentrations (C) of the test strip. The broad source sensitivity (S) for each window is derived using eqn [23] in Section 11.14.4.2.5, where S and C have the same meaning but N is now the fully corrected and stripped count rate in the K, U, or Th window at the nominal survey altitude.

The calculated potassium, uranium, and thorium concentrations can then be determined from field survey measurements by dividing the measured, fully corrected K, U, or Th window count rate (N) by the appropriate K, U, or Th sensitivity (S).

11.14.4.3.7 Airborne gamma ray spectrometer field verification tests

During the course of the survey, tests must be carried out periodically to ensure that the spectrometry system is functioning correctly. In addition to monitoring the energy resolution and energy calibration daily, each day prior to commencing flying operations, a static test must be carried out to ensure overall system sensitivity is maintained by recording an on-ground 'background' or base station spectrum for 5 min followed by a 5 min spectrum with thorium source(s) placed at precisely repeatable locations and the aircraft parked at a precisely repeatable location. The tests will confirm the stability of the overall recorded spectrum.

In addition, a test line near the survey location must be flown at the start and if possible, at the end of each sortie to ensure that the overall sensitivity of the GRS system is maintained constant under operational conditions. This line must be long enough to permit the acquisition of at least 100 s of data while flying at normal survey speed and altitude.

To facilitate monitoring of background radon changes, ideally, daily tests will include flights at survey altitude at the beginning and end of each day over a large body of water. These flights will be at least 500 m from land, with data being acquired for at least 100 s. While always useful, daily over water flights may not be necessary depending on the method of radon correction used. For example, when using an upward detector, sufficient data from over water must be obtained to be able to calibrate the upward detector. Such data need not necessarily be collected before and after every flight, but can be gathered over time as and when possible and convenient. However, daily overwater flights would be a requirement if direct subtraction of radon background derived from overwater flights is the method employed.

11.14.4.4 Data Acquisition with Borehole Logging Systems

11.14.4.4.1 Introduction

Borehole logging systems are used to make measurements inside boreholes (or drill holes) with borehole probes containing sensors and instrumentation, which transmit the data via a cable, to recording instruments on the surface. The logging cable and probe are lowered down the hole from a drum on a winch with appropriate slip rings on its axis to pass the electrical signals from the moving cable to the stationary logging system. The cable passes over a pulley with a depth encoder, which provides the logging system with information on the depth of the probe in the hole as it moves down or up. A borehole log is a record or graph of the sensor output (e.g., gamma ray count rate) plotted on the horizontal axis versus depth on the vertical axis.

Different probe sensors are available, which can measure physical properties of the rocks or soils surrounding the hole, such as their magnetic properties, electrical resistivity, radioactivity, acoustic velocity, density, and porosity. Some probes contain more than one sensor and can produce more than one log from each trip down a hole. The logging systems can be either backpack-portable with a small winch and cable for use in shallow holes or truck-mounted, with large winches carrying cables with length up to several kilometers for logging deep holes such those commonly drilled in hydrocarbon exploration. Logging speed is maintained constant and is usually set at speeds ranging from about 0.3 m min^{-1} to over 10 m min^{-1} depending on the objective of the survey. Some applications require stationary measurements often made stepwise in the hole. Applications range from mineral and hydrocarbon exploration to environmental investigations. An overview of borehole geophysical logging techniques for mineral exploration was given by Killeen (1997a).

The radiometric techniques, which are described here, include two forms of natural gamma ray logging: TC gamma ray logging (also called gross count gamma ray logging) and spectrometric gamma ray logging (also called spectral gamma ray logging). Most commonly, the sensors (detectors) in the probe are NaI (Tl), but CsI (Na) and BGO are also used. The borehole diameter imposes constraints on the maximum diameter of the probe and the size of the detector inside it. In mineral exploration, probes with diameters as small as 32 mm with detector diameters as small as 13 mm are not uncommon. Gross count gamma ray logging was in use for about 20 years

before spectral gamma ray logging was developed, and it was often simply called gamma logging. The term 'gamma logging' is often used somewhat loosely today to refer to 'TC' or 'spectral' gamma ray logging albeit with some ambiguity.

11.14.4.4.2 TC gamma ray and gamma ray spectral logging

TC gamma ray logs refer to counting gamma rays above a certain energy limit using a LLD as described earlier in the section on fundamentals. For most purposes, this energy limit is 400 keV. Early gamma logging was equivalent to using a scintillometer and resulted in the term 'gross count logging.' Today, TC gamma ray logs are usually the TC window log as determined from a gamma ray spectral log. Much of the early use of gross count logging was for determining the grade and thickness of ore zones in uranium exploration and mining using a properly calibrated gamma ray probe. When the detector in the probe passes through a mineralized (uraniferous) zone penetrated by the hole, the recorded gamma ray count rate increases, producing an anomalous peak in the log, which can be analyzed to determine the uranium grade (concentration) and the thickness of the zone (Conaway and Killeen, 1978a,b; Scott, 1963; Scott et al., 1961). For many years, the gamma ray logs were submitted to the US Department of Energy (USDOE), which used the logs to estimate the country's uranium reserves. The USDOE also provided the required uranium model borehole calibration facilities in Grand Junction, Colorado. The first assumption in analyzing the anomaly peaks in the logs was that uranium was the sole source of the gamma rays and that thorium or potassium concentrations in the zones were zero or negligible. This was generally true for uranium deposits in the United States in the early days of uranium exploration. However, new geologic settings for uranium deposits with significant thorium were discovered, such as the pegmatite type at Bancroft and the conglomeratic deposits at Elliot Lake (both in Ontario, Canada). Thus, there was a need for the development of a gamma ray spectral logging system, which could measure Th and K as well as U. The second assumption was that the uranium is in radioactive equilibrium, which was not the case in roll-front sandstone uranium deposits (Hansink, 1976).

In another important application of gross count gamma ray logs, the oil industry used the logs to differentiate between sandstones and shales by their gamma ray activity originating from potassium. The sandstones, which are of interest for their oil potential, had a low K content and showed a low count rate on the log, while the shales, which had a relatively high K content, showed a high gamma ray count rate on the log. This concept worked well until exploration for North Sea oil found that the sandstones there could contain significant amounts of thorium minerals, which caused errors in the sandstone/shale differentiation. The solution was to use gamma ray spectral logging.

The fundamentals of gamma ray spectral logging measurements are very similar to measurements with a portable gamma ray spectrometer as described earlier. Spectral logging uses the same energy windows for the TC, K, U, and Th windows as recommended by the IAEA, the same method of energy calibration and dead time corrections, and the same concerns about stabilization. The stripping ratios and sensitivities for the K, U, and Th windows are similarly determined with a calibration facility, in this case model boreholes instead of calibration pads.

Correction factors are required for both TC and spectral gamma ray logging to account for changes in the borehole conditions from hole to hole. These relate to the hole diameter, whether the hole is air-filled or water-filled (or other fluid, e.g., drilling mud), and whether the hole is cased or open. Methods of data processing are also basically the same for TC logs and for each of the four logs from spectral gamma ray logging.

11.14.4.4.3 Background radiation in borehole measurements

Generally, in the 4π geometry of a borehole gamma ray measurement, there is no background radiation because the detector is entirely inside the sample being analyzed. Two situations, however, are possible in which the background becomes significant. The first occurs in hydrocarbon exploration when drilling muds rich in potassium compounds, such as potassium chloride (KCl) and potassium bicarbonate ($KHCO_3$), have been used. Cox and Raymer (1976) discussed the effect of potassium muds on gamma ray logs and concluded that the contribution of mud radioactivity to the gamma ray measurement can be large. Background corrections are unnecessary for most hard rock mining applications, which use small diameter boreholes and water as a drilling fluid.

The second situation in which the background in a borehole becomes significant occurs especially in uranium exploration logging, when an air-filled borehole, which intersects a uranium-rich zone, is logged after a considerable time has elapsed since drilling. In this case, it is possible that radon, which is a gas, will migrate from radioactive zones to the borehole, creating a broader anomaly on the gamma ray log than if the hole had been logged shortly after drilling. This has been reported as a problem in air-filled model boreholes used for calibration, but the effect can be reduced to an insignificant amount by keeping water in the model holes (IAEA, 1976). The characteristics of radon emanation and corrections for it have been discussed in detail by Scott and Dodd (1960), Austin (1975), and Barretto (1975).

Radon was also reported to be a problem in an environmental study using gamma ray spectral logging of air-filled holes in a landfill site, which contained low-level radioactive waste from a radium refinery (Pflug et al., 1993). In this case, a drilling and logging protocol was developed involving air flushing of the holes before and during logging to minimize the effect of radon.

11.14.4.4.4 Determination of borehole correction factors

To be able to compare results of any type of gamma ray logging from one hole to another hole, a set of standard conditions for logging a hole must be established both for use in calibration models and for use in the field. These conditions are chosen to be those most commonly found in holes encountered in the field, so that the number of corrections will be minimal. Correction factors are the ratio of the gamma ray counts recorded in standard conditions to the counts recorded in nonstandard conditions. For example, calibration factors could be determined in calibration model boreholes for the standard conditions of a hole diameter of 114 mm, air-filled and with no casing, which is the standard at the Grand Junction,

Colorado, calibration facilities. In Canada, the Bells Corners (Ottawa) calibration standard condition is a hole diameter of 60 mm, water-filled and with no casing. Both calibration facilities also have model holes, which can be used to derive the correction factors for logs recorded in nonstandard conditions. For example, there is a large concrete block with holes of several diameters, which can be logged to determine the hole size correction from the difference in count rates measured in each hole. Water correction factors are determined by logging the holes air-filled and water-filled. Similarly, casing correction factors are determined by logging the holes with and without a piece of casing attached to the probe as a sleeve to simulate the effect of a cased hole. Usually, various types (e.g., steel and plastic) of casing sleeves with different thickness are available at the site. The correction factors are applied to logs recorded in the field to adjust the count rates for any differences from the standard conditions that were adopted. Dodd and Eschliman (1972) described how a set of appropriate model holes can be used to provide water-correction factors, casing correction factors, and dead time measurements. They also pointed out for uranium exploration the nonlinear response of total gamma ray intensity from uranium caused by the increasing equivalent atomic number (Z_{eq}) of the rock as the ore grade increases. This 'Z-effect' is due to the high Z uranium attenuating its own gamma rays. Gamma rays of low energies are particularly strongly attenuated when the Z_{eq} increases above values typical of common rocks. As Z_{eq} increases, the proportion of photoelectric effect interactions increases for low-energy gamma rays (**Figure 3**). However, for gamma rays above about 400 keV, there is little change and there should be a linear relation between gamma ray count and the ore grade–thickness product. In gamma ray spectral logging, a TC channel with a 400 keV low-level threshold will provide count rates, which vary nearly linearly with ore grade (for 'infinitely' thick zones).

11.14.4.4.5 Calibration of natural gamma ray logging systems

Uncalibrated gamma ray logs are presented as cps. The cps units are, however, dependent on the detector type and size. This makes it difficult to compare data acquired by different companies using different borehole probes and detectors. The API developed the API test pits (calibration model boreholes) at the University of Houston to standardize gamma ray logging responses from different logging tools into API units. The API gamma ray test pit is a concrete model borehole with zones of high and low radioactivity. The difference in curve deflection on the gamma log between the low and high zones is defined as 200 API units. Gamma ray logs presented in API units are thus standardized from the calibration model borehole so that the logs are independent of the detector characteristics. However, the API unit does not provide information about the radioelements K, U, and Th, which are the source of the measured gamma rays.

To make quantitative measurements of the concentration of the radioelements K, U, and Th by TC gamma ray logging or spectral gamma ray logging, the borehole probes must be calibrated in model boreholes with known concentrations of the radioelements. In the early days of gross count logging for uranium exploration and mining, it was thought that probes could be calibrated for in-hole uranium assays by comparing the uranium assays of drill core to count rates on the gamma ray log. For a large number of different uranium grades and associated count rates, it was hoped that a linear relation between count rate and concentration would be determined. However, the result was a 'buckshot' scatter of points on the graph. It was found that both the ore grade and the thickness of the mineralized zones affected the count rate. Zones with the same ore grade but different thickness produced different gamma ray count rates. This is because the thicker zones occupied a larger percentage of the measurement sample volume and therefore produced a higher count rate than the thinner zones. Only for zones that were infinitely thick (about 1 m), which completely filled the sample volume, would there be a linear relation between count rate and uranium grade.

Quantitative uranium determinations from TC logs were made possible when it was determined that the area beneath the anomaly curve on the log is proportional to the quantity of radioactive material in that part of the borehole (Scott et al., 1961). This is expressed as

$$GT = KA \qquad [24]$$

where G is the average ore grade in % eU (or other units) over a zone, which is T meters thick. K is the calibration factor (sensitivity) determined in model boreholes and A is the area under the anomaly curve in meters $\times$ cps. K has units of % eU $\times$ $(\text{cps})^{-1}$.

Once the K factor is known, it can be used in the field to determine the grade–thickness product of uranium-mineralized zones from the area under the anomaly curve in the log. A gamma ray spectral log is a set of four logs, the TC log, K log, U log, and Th log. Sensitivities for each of the four logs can also be similarly determined in appropriate model boreholes.

11.14.4.4.5.1 Calibration for TC gamma ray logging

A great deal of information is available on TC gamma ray logging and calibration for quantitative measurements in uranium exploration. It assumes that there is negligible associated thorium mineralization; otherwise, a gamma spectral log is required (Conaway and Killeen, 1978a; Rhodes and Mott, 1966; Scott, 1963; Scott et al., 1961).

As mentioned earlier, the calibration of a TC or gross count gamma ray logging system requires the determination of the K factor (or sensitivity) in a model borehole. The grade (G)–thickness (T) product (GT) of a radioactive zone in a model borehole is known from its construction. The area (A) under the anomaly curve of a gamma ray log in the model can be determined from a slow log of the hole. Then,

$$K = GTA^{-1} \qquad [25]$$

Calibration model boreholes are usually constructed of concrete with an 'infinitely' thick ore zone sandwiched between two barren zones as shown in **Figure 10**. A slow gamma ray log of the hole starting in a barren zone will produce a count rate anomaly that rises from the background count in the barren zone as the detector approaches the ore zone and reaches a maximum count rate when the sample volume is entirely inside the zone. It remains at that count rate (i.e., 'flattens out') for several measurements and then begins to decrease when the sample volume starts to include some of the other barren zones. The count rate continues to fall until it reaches background again when the entire sample

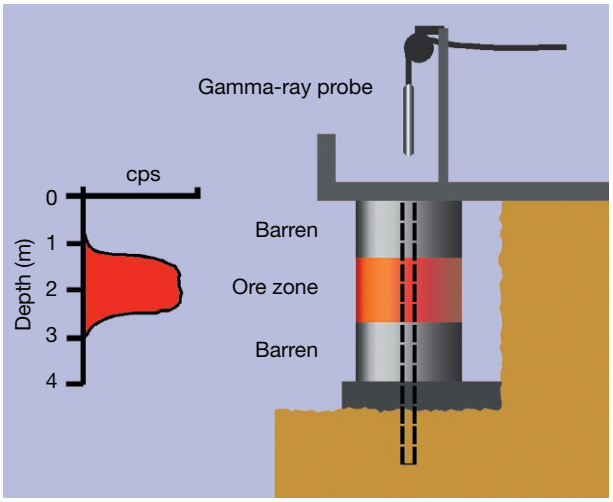


Figure 10 Concept of a calibration model borehole constructed of concrete with an artificial ore zone containing known concentrations of K, U, or Th minerals, between two barren zones. Gamma ray probes are calibrated by analyzing the gamma ray log (left), in counts per second (cps) obtained by logging through the model.

volume is in the second barren zone (Figure 10). The area under the curve is computed by summing the count rate for each sample (measurement) multiplied by the sample interval each measurement represents. For example, at a logging speed of 0.6 m min^{-1}, and a sample time of 1 s, each measurement represents a sample interval of 1 cm (0.01 m) in the hole. Once the area A is computed from background to background in the model, K can be calculated from eqn [25].

However, in the case of an 'infinitely' thick homogeneous zone (and only in that case), the grade is directly proportional to the peak height or average intensity (I) of the flat top of the gamma ray log anomaly in the middle of the ore zone. This is because the sample volume of the detector is completely inside the ore zone for a number of measurements, which constitutes the flat top or plateau of the anomaly. Then,

$$G = KI \qquad [26]$$

The determination of K is thus simplified by making measurements in model holes with 'infinitely' thick ore zones (approximately 1 m thick or greater). Calibration facilities are constructed to make this method possible but the 'area' method can also be used in an infinitely thick ore zone.

For deconvolution or inverse filtering of gamma ray logs, as described by Conaway and Killeen (1978a), and in the data processing section in the succeeding text, the shape constant α (as described in Section 11.14.5.4.1) must be determined in model boreholes. The method is based on taking the natural logarithm of the exponential rise in count rate as the detector approaches the ore zone in the model borehole. The natural logarithm curve shows a linear portion, which has a slope equal to the shape constant α. A method to derive the shape constant α in field boreholes was also described by Conaway (1980b).

11.14.4.4.5.2 Calibration for spectrometric gamma ray logging

Gamma ray spectral logging systems are also calibrated in model holes containing 'ore zones' of thorium, and of potassium, in addition to the usual uranium ore zone found in model holes used to calibrate TC logging systems. The calibration procedure is similar to that described for calibration of surface gamma ray spectrometer systems, which use four calibration pads: K, U, Th, and a blank. For a borehole gamma ray spectrometer, the grade and thickness of the ore and barren zones assigned to the K, U, and Th models are related to the count rate data from a slow gamma ray spectral log in each model hole through the upper barren zone, ore zone, and lower barren zone. A statistically adequate count rate must be obtained over the middle of each ore zone and barren (blank) zone. These count rates and the radioelement analysis (grade) of the K, U, and Th ore zones and the barren zone material are the input data to a regression analysis computer program, which solves for the values of the sensitivities and stripping factors as described in Section 11.14.4.2.4.

Instead of one K factor (sensitivity) as for the TC gamma ray log, there are three sensitivities called 'C factors,' which relate to the area under the curve of the gamma ray log for each radioelement, that is, the K log, U log, and Th log. So instead of eqn [24], we have

$$G_K T_K = C_K A_K \qquad [27]$$

$$G_U T_U = C_U A_U \qquad [28]$$

$$G_{Th} T_{Th} = C_{Th} A_{Th} \qquad [29]$$

where the subscripts K, U, and Th identify the G, T, C, and A for each of the K, U, and Th ore zones. The three C factors are determined in the same way as the K factor for TC gamma ray logs described earlier.

11.14.5 Data Processing

11.14.5.1 Introduction

Data processing for surface, airborne, and borehole radiometric survey techniques, in particular, gamma ray spectrometers, includes all of the computations and corrections that must be applied to the recorded data to ultimately lead to the determination of the concentration of K, U, and Th, either on the ground or in the wall rocks of a borehole. It includes corrections related to the instrument, corrections for background radiation and attenuation of the gamma rays by air or water, application of the stripping ratios, and finally application of the sensitivities to measure K, U, and Th concentrations. Data processing for surface and airborne systems is described first.

11.14.5.2 Data Processing for Surface Systems

11.14.5.2.1 Portable gamma ray spectrometers

Most of the data processing for portable gamma ray spectrometers is related to the application of the factors determined during calibration. For each measurement, the recorded gamma ray count data are processed as follows. Depending on the instrument, the dead time corrections are first applied to the counts and the background is subtracted. (Note that dead time corrections are often done automatically in modern portable spectrometers.) The stripping ratios are then used to strip the counts in the K, U, and Th windows using the stripping

equations [30]–[33] to produce the stripped count rates (N) (Grasty, 1977; IAEA, 1991). Note: there are some typographic errors in IAEA (1991), which have been corrected here:

$$D = 1 - g\gamma - a(\alpha - g\beta) - b(\beta - \alpha\gamma) \quad [30]$$

$$N_K = [n_{Th}(\alpha\gamma - \beta) + n_U(a\beta - \gamma) + n_K(1 - a\alpha)]D^{-1} \quad [31]$$

$$N_U = [n_{Th}(g\beta - \alpha) + n_U(1 - b\beta) + n_K(b\alpha - g)]D^{-1} \quad [32]$$

$$N_{Th} = [n_{Th}(1 - g\gamma) + n_U(b\gamma - a)a + n_K(ag - b)D^{-1}] \quad [33]$$

where α is the Th into U stripping ratio, β is the Th into K stripping ratio, γ is the U into K stripping ratio, a is the U into Th 'reverse' stripping ratio, b is the K into Th reverse stripping ratio, g is the K into U reverse stripping ratio, and n_K, n_U, and n_{Th} are the raw unstripped window counts.

Since b is equal to 0 for the windows described earlier (see Sections 11.14.2.8.1 and 11.14.4.3.3), using the five remaining stripping ratios, the background-corrected count rates in the three windows can be stripped to give the counts in the potassium, uranium, and thorium windows that originate solely from potassium, uranium, and thorium with the following somewhat simplified equations:

$$D = 1 - g\gamma - a(\alpha - g\beta) \quad [30a]$$

$$N_K = [n_{Th}(\alpha\gamma - \beta) + n_U(a\beta - \gamma) + n_K(1 - a\alpha)]D^{-1} \quad [31a]$$

$$N_U = [n_{Th}(g\beta - \alpha) + n_U + n_K g]D^{-1} \quad [32a]$$

$$N_{Th} = [n_{Th}(1 - g\gamma) + n_U a + n_K ag]D^{-1} \quad [33a]$$

The fully corrected, stripped counts are then ready to be converted into radioelement concentrations as % K, ppm eU, and ppm eTh.

This is done by dividing the fully corrected counts for each window by the sensitivity for that window. Thus, the calculated potassium, uranium, and thorium concentrations are determined using the expression for each radioelement:

$$C = NS^{-1} \quad [34]$$

where C is the concentration of radioelement (K%, eU ppm, or eTh ppm), S is the sensitivity for the K, U, or Th window, and N is the count rate for each K, U, or Th window after dead time, background, and stripping corrections.

11.14.5.2.2 Car-borne gamma ray spectrometers

Data processing procedures for car-borne gamma ray spectrometers lie somewhere between the processing procedures for portable spectrometers and for airborne spectrometers. As mentioned earlier, gamma ray spectrometers used in car-borne surveys are often multichannel systems, which can also be used for airborne surveys. Depending on the survey, it may be appropriate to apply some of the spectral noise reduction techniques for multichannel spectra such as the NASVD method or the maximum noise fraction (MNF) method (Grasty and Minty, 1995; Hovgaard and Grasty, 1997) that are used in processing airborne survey data. The methods are applied to the raw measured spectra to improve the signal-to-noise ratio before other processing.

As was the case for portable spectrometers, most of the processing is related to the application of the factors determined during calibration. For each measurement, the dead time corrections are first applied to the counts and the background spectrum is subtracted. The stripping ratios are then applied to strip the counts in the K, U, and Th windows using the stripping equations given in the section on calibration of portable spectrometers. The radioelement concentrations, % K, ppm eU, and ppm eTh, are then computed by applying the appropriate sensitivity determined on the test area, to the fully corrected, stripped counts in the K, U, and Th windows. Because measurements are usually made on roads, regularly spaced survey lines are usually not possible and it is difficult to produce contour maps. The radioelement ratios Th/U, Th/K, and U/K are often computed and plotted as profiles along with TC, % K, ppm eU, and ppm eTh making a profile of seven parameters.

11.14.5.3 Data Processing for Airborne Gamma Ray Spectrometer Systems

11.14.5.3.1 Noise reduction

Before carrying out any of the computation and corrections mentioned earlier, for multichannel airborne gamma spectrometric systems, there are methods that can be applied to the raw spectra to enhance the signal-to-noise ratio. Spectral enhancement techniques such as the NASVD method or MNF method are often used (IAEA, 2003).

Using a technique similar to principal component (PC) analysis and using groups of raw multichannel spectra (say from individual flight lines), the NASVD method computes the different spectral shapes (PCs) making up the measured spectra. Unlike standard PC analysis, NASVD takes into consideration variations due to Poisson counting statistics from low energy channels to high energy channels. Low energy channels have higher counts and therefore lower statistical noise than the high energy channels. NASVD is able to adjust for the varying Poisson noise across the spectrum so that the spectral components reflect the real spectral shapes. The lower-order components strongly contribute to the spectral shape, and the higher-order components contribute less to the spectral shape and have a low signal-to-noise ratio. New multichannel spectra are then reconstructed by combining the lower-order spectral components, excluding components that do not contribute significantly to the shape of the measured multichannel spectra (i.e., they are mostly noise). The reconstructed spectra have a much larger signal-to-noise ratio than the measured multichannel spectra. The noise-adjusted (i.e., reduced noise) spectra are then processed in the normal way. It has been estimated that NASVD can yield an improvement similar to that obtained by increasing the detector volume by a factor of between three and four (Hovgaard and Grasty, 1997).

The second noise reduction method (MNF) also uses a technique similar to PC analysis, when a sample of noise is available. The noise sample is obtained from the survey data by forward-differencing of the spectra along each line (IAEA, 2003). The more general MNF method (Green et al., 1988; Lee et al., 1990) involves two linear transforms: The first transforms the data to unit error variance in each channel with no covariance between channel count rate errors (a noise-whitening transformation); the second is a standard PC transform on the now 'noise-whitened' data. The two transforms are

combined into the MNF transform. In a similar fashion to the NASVD method, the MNF transform sorts data into higher- and lower-order components by decreasing signal-to-noise ratio. The noise is then removed by discarding the higher-order components before doing the inverse transform.

Sorting the raw spectra into clusters on the basis of similarity in spectral shape can improve results from both the NASVD and MNF methods (Minty and McFadden, 1998). Both methods showed similar results when applied to airborne gamma ray spectra (Minty, 2000, 2003).

11.14.5.3.2 AGRS windowing procedure and energy calibration

The count rates in the various spectral windows in the survey data must first be determined. The window positions (channel numbers) in the multichannel spectra require proper energy calibration, which can be done using the measured, recorded survey data itself. The simplest method of carrying out this windowing procedure is to acquire spectra summed over whole survey flight lines or groups of lines and to use the spectra to identify the channels in which the potassium 1460 keV and the thorium 2615 keV photopeaks occur. The potassium and thorium windows are then centered on these channels, and the uranium window is defined in proportion to them with window energy widths as specified earlier.

11.14.5.3.3 Dead time correction

The first step in the reduction sequence for radiometric data is dead time correction using measured dead time (or live time) data. Dead time correction is made to each window using eqn [9a] or [9b] in **Section 11.14.3.3.5**.

Dead time correction should be applied to each window in the downward-looking detector, including the TC window but not the cosmic window, and not to the upward-looking data as these are processed by different circuits. Some modern spectrometer systems with a separate fast ADC for each detector have essentially zero dead time (except in a few exceptional circumstances) making the correction less important than it is with older systems.

11.14.5.3.4 Filtering to prepare for background corrections

Digital filters should be applied to radar altimeter data to smooth sudden jumps that can arise when flying over steep terrain, which causes problems when height-correcting the data. A median filter or other nonlinear filter is suitable to remove spikes. The spectrometer cosmic channel must also be filtered to reduce statistical noise. In this case, an 11–21-point filter (a simple running average) is recommended. To calculate radon background from the upward-looking detector data as described later, heavily filtered (running average) uranium upward, uranium downward, and thorium downward data are needed.

11.14.5.3.5 Cosmic and aircraft background

As mentioned in the data acquisition section earlier (**Section 11.14.4.3.4**), the cosmic window data are obtained during calibration flights at 500 m intervals between 1500 and 3500 m ASL. For each spectral window, the cosmic and aircraft background is determined using the following equation (IAEA, 1991, 2003):

$$N = a + bC \quad [35]$$

where N is the combined cosmic and aircraft background in each spectral window, a is the aircraft background in the window, C is the cosmic channel count, and b is the cosmic stripping factor for the window.

Using a linear regression of cosmic window count on each of the K, U, and Th window counts for each window, the slope of the line (b) is the cosmic window stripping factor and the intercept of the line (a) is the aircraft background.

The expressions are evaluated for each window at each data point using the filtered cosmic channel data, and the results are subtracted from the window data. This can also be done for every channel in the spectrum instead of for the windows only (IAEA, 2003). Then, the entire aircraft and cosmic background spectrum can be subtracted from the measured spectrum at each data point (sample interval).

11.14.5.3.6 Radon background

There are four methods that can be used for radon background subtraction. Simple direct subtraction estimated from flights over bodies of water close to the survey area (e.g., calibration range as described in **Section 11.14.4.3.5** at Breckenridge) can be employed.

The other three are the spectral-ratio method, the full-spectrum method, and the upward-looking detector method. Of these, the first two methods use the shape of the gamma ray spectrum to estimate the atmospheric radon concentration. These methods cannot be used in the northern hemisphere where there is significant ^{137}Cs contamination caused by fallout from atomic weapon tests in the 1950s and nuclear reactor accidents (e.g., Chernobyl). The third method will be described first, followed by the spectral-ratio method and the full-spectrum method.

11.14.5.3.6.1 The upward-looking detector method

Radon in the air is a problem because one of its decay products is ^{214}Bi, which is used for measuring uranium. An upward-looking detector placed on top of the main detectors, which shield it from the radiation from the ground, can be used to measure the radon background by counting in an 'upward' uranium window. Determination of the constants necessary for the correction of background due to radon, using upward detectors, requires several steps. The procedure outlined in IAEA (1991) is generally correct, but more recent studies have refined the process. First, the contributions of atmospheric radon to the various spectral windows are determined with a series of test flights over water. The count rates in the K, U, Th, and TC windows will include aircraft and cosmic ray background as well as radon background. The aircraft and cosmic ray background can be removed leaving only the radon background (^{214}Bi) count rates in the windows. The series of flights at different times will show count rate changes in all of the windows due to variations in radon background only and the TC, K, Th, and both upward- and downward-looking U window count rates will vary linearly with each other. Linear regression of counts in the downward U window with the counts in the TC, K, Th and upward U windows will define their relationships. The coefficients (slopes of the regression

lines) are similar to stripping ratios for the radon contribution to the various windows. Graphs of the regression lines are shown in an example in IAEA (1991).

The response of the upward-looking detector (uranium window) to radiation from the ground (u_g) is determined next. The ground count rate in the upward detector is related to the ground count rate in the downward detector by (IAEA, 1991)

$$u_g = a_1 U_g + a_2 T_g \qquad [36]$$

where u_g, U_g, and T_g are the window counts originating from uranium and thorium on the ground and a_1 and a_2 are coefficients to be determined.

The more recent procedure recommended by Grasty and Hovgaard (1996) summarized in the succeeding text is believed to be more reliable than that in IAEA (1991) and in addition can be used when bodies of water are unavailable.

In view of the high correlation between radiation in the uranium and thorium windows, it is better to assume that the upward response originating from the ground can be correlated to either counts in the thorium window or counts in the uranium window. This correlation is usually high, but not always. The method is reliable where the correlation can be demonstrated. This is equivalent to assuming that either a_1 or a_2 is equal to 0. Solving for a_1 or a_2 is accomplished by subtracting measurements for the upward window and the uranium window (or thorium window) at approximately 30 s intervals to find a set of differences. The TC window will be used to determine whether the radioactivity is increasing or decreasing. It is necessary to first subtract the counts in the uranium window (or thorium window) from the TC to reduce bias of the final result. If the TC window indicates that the radioactivity is decreasing, the sign of both the upward and downward differences must be reversed. The value of the constant (either a_1 or a_2) is then simply the ratio of the sum of the adjusted differences in the upward window divided by the sum of the adjusted differences in the downward uranium (or thorium) window. The expression for the radon component in the downward uranium window is given by (IAEA, 1991)

$$U_r = (u - a_1 U - a_2 T + a_2 b_T - b_u)(a_u - a_1 - a_2 a_T)^{-1} \qquad [37]$$

where U_r is the radon background detected in the downward U window, u is the measured count in the upward uranium window, U is the measured count in the downward uranium window, T is the measured count in the downward thorium window, a_1, a_2, a_u, and a_T are proportionality factors, and b_u and b_T are backgrounds, which are constants determined experimentally from multiple overwater flights.

Using the values for a_1 or a_2 determined earlier in this equation will result in a reasonable estimate of U_r, which will permit the other windows to be corrected for radon.

The measured count rates u, U, and T must first be corrected for cosmic and aircraft background. The radon counts in the TC, potassium, and thorium windows can be calculated from U_r using eqns 4.10, 4.11, and 4.12 given in IAEA (1991).

Because of the low count rate in the upward uranium window, the window data must be filtered considerably to reduce statistical noise. For a system with two upward-looking detectors of volume 8.4 l, a 200-point running average should be suitable.

Test data flown at survey altitude during the daily verification tests described earlier (**Section 11.14.4.3.7**) are used to monitor changes in atmospheric background and calibrate the upward- and downward-looking detector systems using a procedure recommended and described by the IAEA (1991). Variations in the uranium window can be partly due to radon but also due to soil moisture variations or small changes in the flying height or flight path. Variations due to soil moisture and flight path errors in some cases can be overcome by a simple normalization procedure based on the count in the thorium window. The procedure assumes that a given percentage change in thorium count from the ground will correspond to the same percentage change in the uranium counts from the ground. If this assumption is valid, then the procedure is as follows. First, the average thorium count rate for the tests during the entire survey period is found. Then, for each test, the uranium count rate is multiplied by the average thorium count, divided by the thorium count for that flight. Changes from flight to flight in the resulting normalized uranium count are then due to variations in radon, and corrections can be determined for each flight. However, the response of uranium to varying soil moisture can be inconsistent depending on the actual makeup of the soil. Caution must be employed because there are situations when the uranium count increases but the thorium count is constant or decreases. In general, the thorium count can be used to monitor for periods with stable soil conditions. It is preferable that these data be used rather than attempting to use the thorium count rate to correct the data during periods when soil conditions are changing.

11.14.5.3.6.2 The spectral-ratio method and full-spectrum method

The spectral-ratio and full-spectrum methods use multichannel spectra, while the upward-looking detector method uses window data. The spectral-ratio method is based on the ratio of a low-energy peak to a high-energy peak in the uranium gamma ray spectrum. The height of the peaks above the Compton continuum is used in the method, not the full height. If there is no atmospheric radon, this ratio will be relatively constant as the uranium concentration on the ground changes. However, if there is radon in the air, the low-energy gamma rays from ^{214}Bi at 609 keV from the atmospheric radon will not be attenuated as much as the high-energy gammas from ^{214}Bi at 1760 keV from ground radiation, and the ratio will change. In the spectral-ratio method, a low-energy 'radon' window (L) is set on the 609 keV peak, in addition to the K, U (1760 keV), and Th windows. Minty (1998) showed how the ratio could be used to estimate the atmospheric radon in the measured gamma ray spectrum from the radon contribution to the low-energy radon window (L_r). It involves recording a pure radon spectrum from calibration flights at survey height over water when there is radon in the air. Then, L_r is used to multiply the pure radon spectrum channel by channel for the multichannel to be used for the background subtraction.

The full-spectrum method (Minty, 1992) is based on the assumption that a measured spectrum can be considered as the sum of the four component spectra of K, U, Th, and radon. The radon spectrum is determined as described in the preceding text, and Minty (1998) described a method to simulate the K, U, and Th component spectra with calibration experiments on the ground. The four component spectra are then fitted to

the measured spectrum using the least squares method and the K, U, and Th window counts can be determined.

11.14.5.3.7 Calculation of effective height above ground level

The filtered radar altimeter data will be used to adjust the stripping ratios for altitude and to carry out attenuation corrections. They are then converted to effective height (h_e) at STP by the expression (IAEA, 2003)

$$h_e = h \times 273.15 \times (T + 273.15)^{-1} \times P \times (101.3)^{-1} \quad [38]$$

where h is the observed radar altitude, T is the measured air temperature in °C, and P is the barometric pressure in kPa. (Note: STP is $T = 273.15$ °K, $P = 101.325$ kPa.)

If necessary, the pressure (P) can be estimated from the barometric (or GPS) altitude using the expression (Mechtly, 1973)

$$P = P_0 e^{(H/8581)} \quad [39]$$

where H is the barometric (or GPS) altitude in meters and P_0 is the barometric pressure (at sea level) in kPa.

11.14.5.3.8 Stripping

The stripping ratios, α, β, γ, a, b, and g, are determined over calibration pads in a similar procedure to that for portable gamma ray spectrometers described earlier (see Section 11.14.4.2.4.1), following IAEA-recommended procedures. However, the principal ratios α, β, and γ vary with effective altitude (see Section 11.14.5.3.7) above the ground and should be adjusted before stripping is carried out. The increase in stripping ratio per meter is (IAEA, 1991)

$$\alpha = \alpha_0 + 0.00049 \times h_e \quad [40]$$

$$\beta = \beta_0 + 0.00065 \times h_e \quad [41]$$

$$\gamma = \gamma_0 + 0.00069 \times h_e \quad [42]$$

where α_0, β_0, and γ_0 are the original values before adjusting.

The adjusted stripping ratios are now used to strip the counts in the K, U, and Th windows using the stripping eqns [30]–[33] given in the portable spectrometer section earlier (Section 11.14.5.2.1). The background-corrected count rates in the three windows can be stripped to give the counts in the potassium, uranium, and thorium windows that originate solely from potassium, uranium, and thorium.

11.14.5.3.9 Attenuation correction

The attenuation coefficients are used to correct the counts for variations in altitude from the planned survey altitude. The background-corrected TC and stripped count rates vary exponentially with aircraft altitude. Consequently, the measured count rate is related to the count rate at the nominal survey altitude by the equation

$$N_s = N_m e^{-\mu(h_0 - h)} \quad [43]$$

where N_s is the 'survey' count rate normalized to the nominal (planned) survey altitude, h_0, N_m is the 'measured' background-corrected, stripped count rate at STP effective (or equivalent) height h, and Φ is the attenuation coefficient for that window. Typical values of Φ for TC, K, U, and Th are -0.0067, -0.0082, -0.0084, and -0.0066m^{-1}, respectively, at STP (IAEA, 1991).

11.14.5.3.10 Computation of apparent radioelement concentrations on the ground

The fully corrected count rate data are used to estimate the concentrations in the ground of each of the three radioelements, potassium, uranium, and thorium. The procedure determines the concentrations that would give the observed count rates, if uniformly distributed in an infinite horizontal slab source, that is, a broad source. Because the U and Th windows actually measure ^{214}Bi and ^{208}Tl, respectively, the calculation implicitly assumes radioactive equilibrium in the U and Th decay series. The U and Th concentrations are therefore expressed as equivalent concentrations, eU and eTh.

The calculated potassium, uranium, and thorium concentrations are determined using the same methodology outlined for portable gamma ray spectrometers (Section 11.14.5.2.1) and eqn [34]. The only difference is that the airborne count rate N is also corrected for attenuation.

Some government surveys require an estimate of the natural air absorbed dose rate (NADR) from geologic sources to be made from the apparent concentrations, K%, eU ppm, and eTh ppm, using the expression (IAEA, 2003)

$$\text{NADR} = 13.078 \times \text{K} + 5.675 \times \text{eU} + 2.494 \times \text{eTh} \quad (\text{units} = \text{nGy h}^{-1}(\text{nano-Gray per hour})) \quad [44]$$

11.14.5.3.11 Calculation of radioelement ratios

The ratios of the three radioelements (eU/eTh, eU/K, and eTh/K) are frequently plotted as profiles or contour maps. Due to statistical uncertainties in the individual radioelement measurements, care should be taken in the calculation of these ratios. The acceptable method of determining ratios is as follows:

1. Typically, neglect any data points where the potassium concentration is $< 0.25\%$ as these measurements are likely to be over water.
2. Progressively, sum the element concentrations of adjacent points on either side of the data point until the total accumulated concentrations of both numerator and denominator exceed a threshold value. This threshold is normally set to be equivalent to at least 100 counts for both the numerator and denominator. This threshold will be obtained from the experimentally determined sensitivities.
3. Calculate the ratios using the accumulated sums.

With this method, the errors associated with the calculated ratios will be similar for all data points. Additional useful information on airborne gamma ray spectrometric surveying, which complements all of the previous descriptions of AGRS, is given by Shives et al. (1995), Dickson and Scott (1997), Horsfall (1997), Minty (1997), Minty et al. (1997), Wilford et al. (1997), and IAEA (2013).

11.14.5.4 Data Processing for Borehole Systems

11.14.5.4.1 TC gamma ray logs

Data processing includes the application of all corrections up to the point where the sensitivity (K factor) determined in model boreholes can be used to convert counts to the

grade–thickness product. These include dead time corrections, water and hole size corrections, and casing corrections. The objective of TC gamma ray logging in uranium exploration and mining is first to locate zones of anomalous count rate, which represent uranium mineralization. The anomaly on the gamma log is then analyzed by first determining the area under the anomaly curve from background to background, above and below it. The area is computed by summing the count rate (cps) for each measurement multiplied by the sample interval each measurement represents (IAEA, 1986). Once the area A (in meters $\times$ cps) is computed from the count rate anomaly, the grade–thickness product of the uranium-mineralized zones is determined by multiplying the area by the K factor, which has units of % eU $\times$ (cps) $^{-1}$ (or other grade units), that is, using eqn [24] in **Section 11.14.4.4.5** (the expression $GT = KA$). The GT (grade–thickness) has units of % eU meters (Conaway and Killeen, 1980). The GT information from a number of holes can then be used to calculate the ore tonnage for mining purposes.

The grade–thickness information is useful but it does not provide the actual grade or the thickness of the zone. If the thickness of the zone can be estimated from the shape of the anomaly, then the average grade over that estimated thickness is given by

$$G = K\,A\,T^{-1} \qquad [45]$$

Early attempts to determine the thickness from the anomaly, such as defining it as the full width half maximum, were shown to be erroneous (except for infinitely thick zones), sometimes by as much as 200% error (Conaway and Killeen, 1980). Two methods have been developed to provide more details on the grade and thickness from gamma ray logs recorded in the field. An iterative technique was described by Scott (1962, 1963) and Scott et al. (1961), and a deconvolution or inverse filtering technique was described by Conaway and Killeen (1978a). Both techniques require that the shape of the anomaly produced by an ore zone in a model hole be accurately determined. This shape defines the response of the gamma ray logging system and is used as the basis for the iteration or the inverse filtering (Conaway and Killeen, 1978b; Czubek, 1971, 1972).

The iterative technique considers that any complex anomaly as encountered in field logging can be considered to be the sum of individual anomalies from a number of adjacent thin zones, each, say, 5 cm thick, representing the sample interval of the log. This is known as the principle of superposition. Any anomaly can be broken down into a series of adjacent 5 cm thick zones with different ore grades. The shape of the basic 5 cm thick anomaly is determined by gamma ray logging in calibration models. To analyze the complex field anomaly, it is divided into 5 cm thick layers from top to bottom of the anomaly, and a 'first guess' of the grade in each layer is made, based on the relative count rate in the anomaly at each layer. For each of the 5 cm layers, the anomaly for its particular grade is computed using the basic 5 cm anomaly shape. The count rate contribution in all of the sample intervals, from all of the anomalies, is summed for each layer, which produces a synthetic anomaly (Conaway and Killeen, 1980). The difference in count rates between the recorded anomaly and the synthetic anomaly is compared for each layer. The grades for the layers are then adjusted up or down appropriately, and all of the 5 cm layer anomalies are recomputed using the adjusted grades to produce a new synthetic anomaly. This is repeated over and over until the difference is acceptably small. The result is an explanation of the grade of the zone in terms of a number of adjacent 5 cm layers, each with its own grade. The iterative technique (essentially the fitting of a synthetic anomaly) must be used in postprocessing of the recorded gamma ray logs. Killeen et al. (1978) suggested an inverse filter can be used in nearly real time to produce a deconvolved gamma ray log wherein the gamma ray log scale has been converted to grade rather than count rate, and this has since been implemented in several logging systems.

The inverse filter or deconvolution method is based on the application of digital time series analysis (Conaway and Killeen, 1978a,b). It considers the 'geologic impulse response' (GIR) to be the gamma ray log as measured by an ideal point detector as it logs past an infinitesimally thin layer of radioactive material. The anomaly measured across the infinitesimally thin layer has the shape of an exponential function, which rises in count rate from zero to a maximum point over the zone and then falls off exponentially to zero when the sample volume no longer contains any gamma rays detected from the zone (**Figure 11**).

The GIR is a result of the sample volume of the detector moving toward the thin layer, causing the count rate to start to increase when the first gamma rays are detected, even before the detector reaches the zone. The equation for the GIR is a double-sided function:

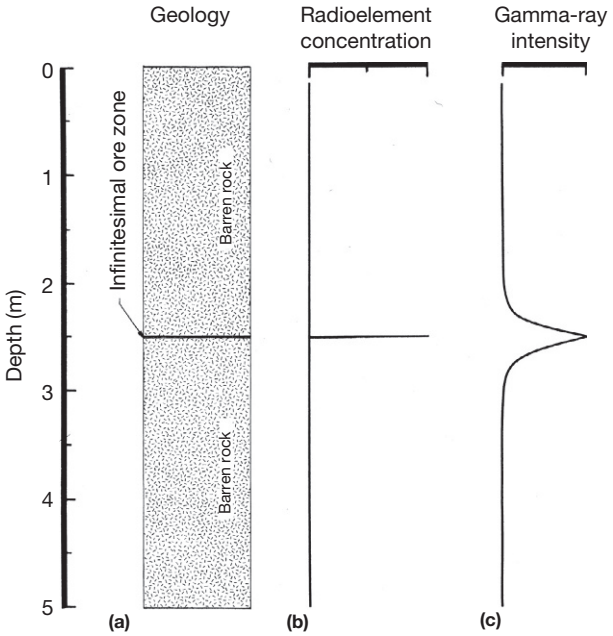


Figure 11 Concept of the GIR: (a) a geologic column, showing an infinitesimally thin radioactive ore zone in homogeneous rock; (b) plot of the radioelement concentration with depth for (a); (c) the 'ideal' gamma ray log response produced by a point detector logging through the infinitely thin ore zone. This is the geologic impulse response or GIR. Reproduced from Conaway JG and Killeen PG (1978) Computer processing of gamma ray logs: Iteration and inverse filtering. In: *Current Research, Part B*, pp. 83–88. Ottawa, ON: Geological Survey of Canada, Paper 78-IB.

$$\varphi(z) = (\alpha 2^{-1})e^{(-\alpha^{|z|})} \quad [46]$$

where φ is the response function, z is depth, and α is a constant. The constant α determines the shape of the GIR, which depends on factors such as density. The true radioelement distribution in the hole is zero everywhere except for the infinitely thin radioactive layer, and this can be called the radiometric assay log (RA log). Instead of the gamma ray log showing this true radioelement distribution, it shows the count rate rising and then falling as the detector passes through the zone. However, the gamma ray log can be deconvolved using an inverse filter or deconvolution operator to produce the RA log (true radioelement distribution), which is what is wanted. The inverse filter is derived for the GIR equation as described by Conaway and Killeen (1978a) and then applied to any gamma ray log to remove the effects of the GIR. The K factor or sensitivity of the logging system and the sampling interval can be incorporated into the three coefficients of the digital inverse operator. Other factors that affect the GIR such as detector length are also described by Conaway and Killeen (1978a), as well as a method to derive the shape constant α in model boreholes. A method for field boreholes was given by Conaway (1980b). Other parameters related to this method were studied such as thin zone corrections and the effect of dipping beds (Conaway, 1978, 1979a, 1981), computer processing of gamma ray logs and inverse filters (Conaway, 1979b, 1980a), and uranium concentrations (Conaway, 1980c). The application of spatial deconvolution to gamma ray logging was reviewed by Conaway (1986). Methods for optimization of gamma ray logging for uranium were given by Conaway et al. (1980). Iteration and inverse filtering were compared and shown to be equivalent for an infinite number of iterations (Conaway and Killeen, 1978b), and Bristow et al. (1984) presented a case study on the application of inverse filters to gamma ray logs. An overview of much of the aforementioned material on gamma ray logging and interpretation was given by Killeen (1982).

11.14.5.4.2 Gamma ray spectral logs

Data processing for gamma ray spectral logs follows much the same procedures as for TC logs in terms of applying corrections for dead time, borehole fluid (e.g., water), hole size, and casing. The primary difference is the application of stripping ratios determined in K, U, and Th model boreholes as part of the calibration procedure. The counts in the K, U, and Th windows are stripped using the equations given earlier in the section on portable gamma ray spectrometers (**Section 11.14.5.2.1**). After stripping and the appropriate corrections have been made, the gamma ray spectral logs (TC log, K log, U log, and Th log) can be processed in the same way as the TC (gross count) log from a nonspectral system as described previously. There are three C factors, one for each of the K, U, and Th logs, determined during calibration, which are applied in the same way as the K factor is used in the relation $GT = KA$. The three logs can be processed to determine each of the % K, ppm eU, and ppm eTh grade–thickness products from the measured areas (A) under the anomaly curves of the K, U, and Th logs, respectively. The iterative technique and the inverse filter method described in the preceding text can both be applied to the three spectral logs to obtain more detailed information about the distribution of the radioelements in the borehole wall (Conaway and Killeen, 1980).

The processing of spectral gamma ray logs is illustrated in **Figure 12**, which shows the TC log with both the raw K, U, and Th logs (before stripping) and the stripped logs converted to radioelement concentrations. The TC log shows anomalies in zones near the top and bottom of the borehole. The raw spectral logs also show anomalously high count rates in a zone near the top of the borehole and in a zone near the bottom of the borehole. The anomalies near the top of the borehole are due to high potassium content as there is no indication of anomalous gamma ray activity in the raw U and Th logs. The anomaly near the bottom of the borehole shows anomalously high gamma ray count rates in all the three raw logs, suggesting that the gamma ray activity is from all three radioelements. However, after stripping and converting the

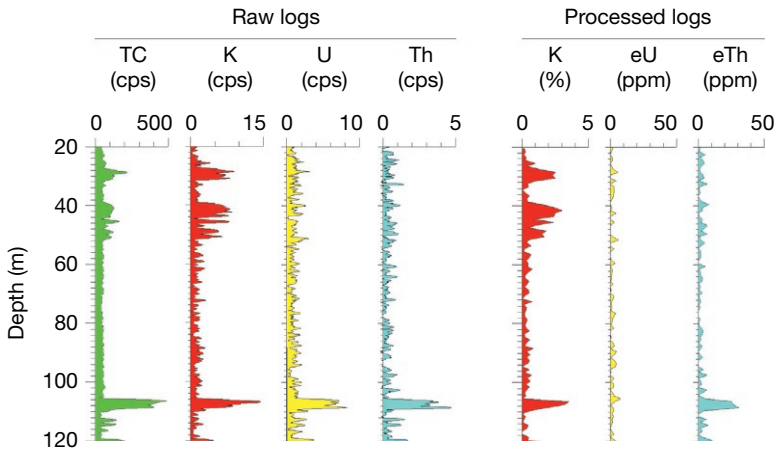


Figure 12 Illustration of the processing of spectral gamma ray logs. The total count log and the raw K, U, and Th logs (before stripping) are shown on the left, and the fully processed, stripped logs converted to radioelement concentrations are on the right. The stripped logs show that there is potassium and thorium but no uranium in the lower zone and that the high count rate in the raw uranium log was caused by gamma rays from thorium. Reproduced from Ford KL, Keating P, and Thomas MD (2007) Overview of geophysical signatures associated with Canadian ore deposits. In Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 939–970. Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

data into radioelement concentrations as described earlier, it is obvious from the stripped logs that there is potassium and thorium but no uranium in the lower zone and that the high count rate in the raw uranium log was caused by gamma rays from thorium.

11.14.6 Data Presentation and Visualization

11.14.6.1 Introduction

Fully processed GRS data (airborne, ground, and borehole), collected with properly calibrated systems, can be displayed in a variety of ways that enhance interpretation. Traditionally, AGRS data have been displayed as single or multi-variable (stacked) profiles for flight line data and as black and white contour maps or color interval maps for gridded data combined with topographic information. With increased usage of geographic information systems (GIS) and image processing techniques, the visual enhancement of AGRS data, by combining it with other remotely sensed data, has become much more common (Ford et al., 2008a,b; Milligan and Gunn, 1997; Schetselaar et al., 2007). Despite these advancements, the traditional stacked-profile and color interval presentations are still in common use today and provide an important visual interpretation tool. The standard practice for borehole GRS (BGRS) data is to plot either TC activity or radioelement content on the horizontal axis with the depth axis plotted vertically. Common data presentation and visualization methods for AGRS and GGRS data are described first followed by BGRS methods.

11.14.6.2 Variables and Standard Survey Products for AGRS Data

11.14.6.2.1 Digital data

For all modern AGRS surveys, the digital data are considered to be the primary and principal product of the survey and must be of the highest quality and error-free. Modern standards for digital data delivery include all raw and final processed line and gridded geophysical data, as well as positional and environmental data in a user-compatible format with the AGRS data commonly at 1 s intervals. This digital data product should include data collected from all calibration and verification tests including daily spectrometer test line and background data, daily thorium source checks, daily or weekly ^{137}Cs energy calibration and resolution tests, and results of all pre- and post-sortie test flights as described in **Section 11.14.4**. The final digital dataset should also include results from any intermediate steps that are not part of regular correction procedures. This may include any nonstandard, manually introduced corrections. Overwater values should not be deleted from a dataset or set to zero as such values contain useful data-quality information. Due to the statistical nature of the GRS data, 'negative' concentrations are to be expected in low average count rate areas. If background corrections are properly determined and applied correctly, then radioactive element concentrations should yield a mean concentration of zero over large bodies of water. The digital data should include the raw spectra recorded in real time and spectra reconstructed using spectral-enhancement methods, such as NASVD or by MNF, both of which improve the signal-to-noise ratio compared to recorded spectra. See **Section 11.14.5** for descriptions of noise reduction techniques. These spectral-enhancement techniques impose rigorous constraints on the means of measuring raw spectra (Grasty and Minty, 1995). The use of these techniques will not change the average responses for an area but will reduce the noise in individual readings. See previous sections of this chapter for a more comprehensive description of AGRS fundamentals, calibration, and data processing techniques.

11.14.6.2.2 Gridding and color interval maps

AGRS data are usually represented by four measured variables, potassium, eU, eTh, and total air absorbed dose rate (TADR) (nGy h^{-1}). TADR is the estimated total air absorbed dose rate computed from the measured TC window corrected to ground level. Five derived products are also required to fully present the AGRS data. These include the three radioelement ratios (eU/eTh, eU/K (ppm/%), and eTh/K (ppm/%)). A three-component or ternary radioelement map similar to that developed by Broome et al. (1987) and the NADR (nGy h^{-1}) have also become common presentations. NADR is the natural air absorbed dose rate at ground level computed from the measured potassium, eU and eTh concentrations. Ford et al. (2008a,b) presented examples of standard color interval and ternary radioelement maps compiled from both reconnaissance (5000 m line spacing) and regional (1000 m line spacing) AGRS survey data. Most map products require the data to be interpolated onto a regular grid. Many of the standard gridding algorithms are unsuited for direct use with AGRS data because of the inherent statistical variations where sampling density along a flight line can be significantly greater than that across lines (IAEA, 2003). A suitable gridding algorithm is one that takes the average of all data points lying within a circular or elliptical area, inversely weighted for distance from the grid point. Many of the common geophysical processing packages available offer a selection of suitable gridding algorithms including bidirectional, minimum curvature, moving average, and kriging, which should only be applied to AGRS data if appropriate smoothing or filtering has been carried out prior to gridding. A more comprehensive explanation of these four commonly used methods can be found in IAEA (2003).

The three-component or ternary radioelement map similar to those shown in **Figures 13** and **14** is an effective method to display variations in total radioactivity and the relative abundances of the three radioelements. Areas of the image with the same color will have similar ratios of K, eU, and eTh, and the intensity of that color is a measure of the total radioactivity. This allows the map to represent the radioelement distribution better than any of the other single variable or ratio maps. A ternary image provides an initial overview of the distribution of radioactivity; however, it does not replace the more detailed, quantitative information available on the other seven maps (TADR or NADR, K, eU, eTh, eU/eTh, eU/K, and eTh/K).

Often, depending on the complexity of the geology, subtle variations in K, eU, and/or eTh may not be readily apparent on a single variable map. For this reason, the proper display and interpretation of GRS data require the examination of all of the measured variables and associated derived products. Ratio maps can enhance or reinforce subtle variations in the

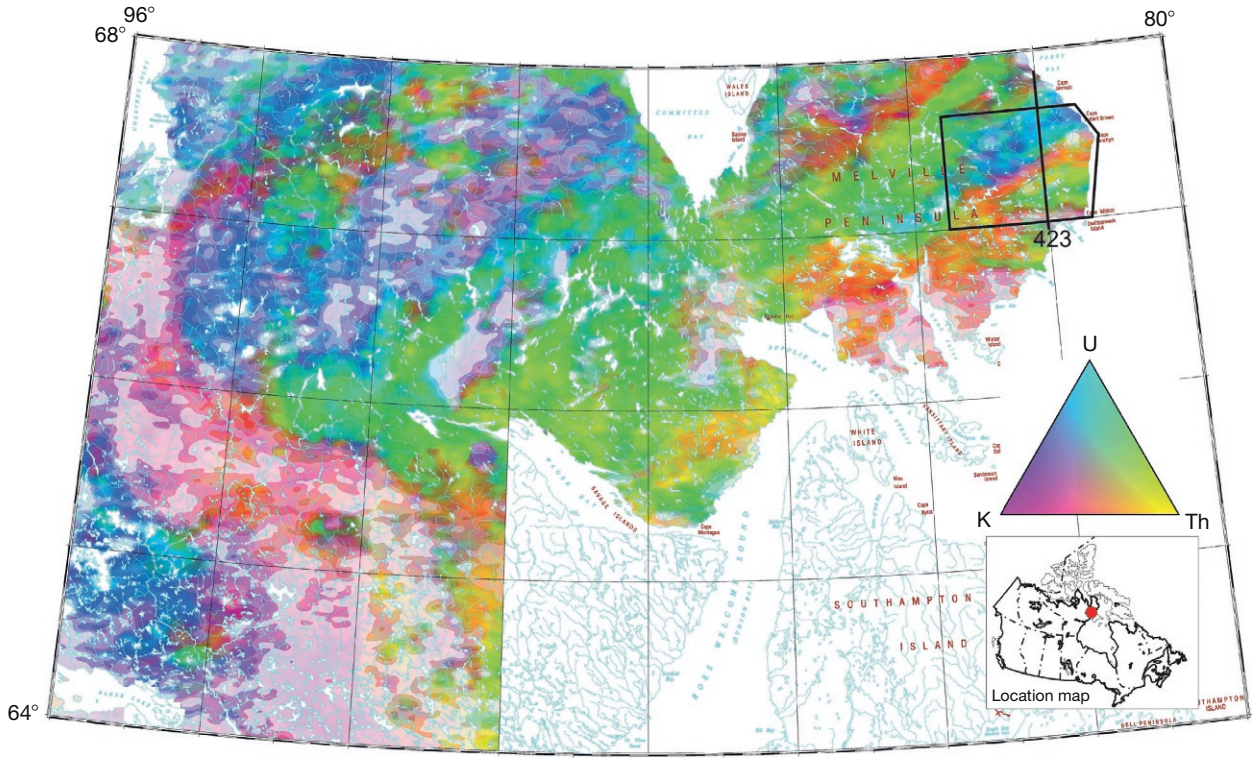


Figure 13 Ternary radioelement image for the Quoich River area, Nunavut. The Melville Peninsula area is outlined in black. The location of flight line 423, used in **Figure 16**, is also indicated. This image was produced from 5000 m line spaced AGRS data. Reproduced from Holman PB, Carson JM, Ford KL, Grant JA, and Shives RBK (2001) Airborne gamma ray spectrometry compilation series, Quoich River, Nunavut. Geological Survey of Canada, Open File 4005, scale 1:1 000 000.

measured variables. This may be particularly important when trying to map subtle cryptic compositional variations in a suite of granitic intrusions (Ford, 1993; Ford and O'Reilly, 1985) or when dealing with varying intensities of alteration associated with a mineralizing process (Shives et al., 1997). Grids of the radioelement ratios should be produced directly from the gridded concentration data. An acceptable method would see grid cells that have a potassium value less than that, which reflects an average overwater value (e.g., 0.25%) be set to undefined. Cells that have potassium values greater than this threshold can be used to compute ratios. The grids for the numerator and denominator should be ring-searched to ensure both exceed the equivalent of a 100-count threshold. Only cells that have a potassium concentration $>0.25\%$ will be included in this accumulation process.

Like other airborne geophysical data, line spacing and spatial resolution play a significant role in determining how much geologic detail can be extracted from the imagery, either in a visual sense or through the use of computer-processing techniques. For AGRS data, this is clearly illustrated in **Figure 15**, which shows gridded, ternary images of an area in southwestern Nova Scotia, Canada, at three different line spacings (Ford et al., 2008a,b). It is interesting to note in these images that the same basic geologic information, for example, major geologic features in a granite-dominated terrain, is also readily visible at wider line spacing (lower resolution) as compared to the higher-resolution surveys. In fact, the overall correlation between the images acquired at different line spacings is statistically similar (**Table 6**). The difference is in the detail. In fact, details of rapidly varying surface cover conditions, such as soil moisture and surface water, are often suppressed on lower-resolution data in which much of the response is from regional bedrock patterns. Survey objectives, budget, and the geologic complexity of a survey area are some factors that must be taken into consideration when deciding on the appropriate flight line spacing. If the objective is to define regional radioelement signatures to support reconnaissance mapping applications, then lower-resolution line spacing may be suitable as illustrated in the example from granite-dominated terrains (**Figure 15**). Such wide line spacing will likely not be suitable in sedimentary-dominated terrains or where more detailed information is required to support planned or future detailed mapping or mineral exploration activities.

11.14.6.2.3 Line data presentation

Gridding AGRS data is required to produce color interval maps or images that are useful for regional mapping applications and to visually represent the geographic variations of a single variable or ratio. However, a multivariable stacked-profile display is an essential tool for the visual analysis of all variables at each sample point including other geophysical sensor or environmental data. A stacked-profile presentation format permits the viewing of point-to-point data for all variables measured and derived. Ratios should be determined using radioelement concentrations and not corrected counts. See **Section 11.14.5** for an acceptable method of determining ratios for line data

presentation. With this method, the errors associated with the calculated ratios will be similar for all data points. Gridding AGRS data, particularly if the data were collected with coarse flight line spacing, requires a significant amount of interpolation, which will result in the loss of some detailed point-to-point information. Many of the small-scale variations that may be useful for mineral exploration applications (e.g., identification of small anomalies) are averaged out in the interpolation process. This supports the use of both gridded and profile data from regional to more detailed mapping studies. Viewing of the point-to-point or line data in a stacked-profile format requires only a minimum amount of along-line filtering, thereby preserving more of the detailed point-to-point information. Examples of multivariable stacked profiles are shown in **Figure 16(a)** and **16(b)**.

11.14.6.2.4 Cross plots for surface and airborne gamma ray spectrometry data

To properly interpret or validate the geologic information provided by an AGRS survey, a fully calibrated GGRS is often used. Modern GGRS have become smaller and very easy to operate. For follow-up of an AGRS survey or to support regional mapping or mineral exploration programs, a GGRS can provide real-time in situ geochemical information that can be used to improve bedrock or surficial mapping and to improve sampling efficiencies by helping to guide subsequent sampling. Data collected with a GGRS are often presented as a cross plot (XY graph) that can be used to show interrelationships between the three measured radioelements, K (%), eU (ppm), and eTh (ppm). Such graphs are used to show radioelement clusters that support lithologic mapping by defining characteristic signatures and by indicating trends that deviate from the normal radioelement signatures (**Figure 17(b)** and **17(c)**). These deviations may be associated with large-scale cryptic compositional variations within intrusive rocks that can be tied to other lithogeochemical variations (Ford, 1993) or may indicate alteration associated with certain types of mineralization (Shives et al., 1997). Such applications demonstrate the effective use of radioelement variations mapped by AGRS or GGRS as indicator or pathfinder elements for other compositional variations. With the increasing importance of remote predictive mapping (RPM) studies, the data collected during a GGRS survey can be used to support and validate classification and cluster analysis of AGRS data (Harris et al., 2008).

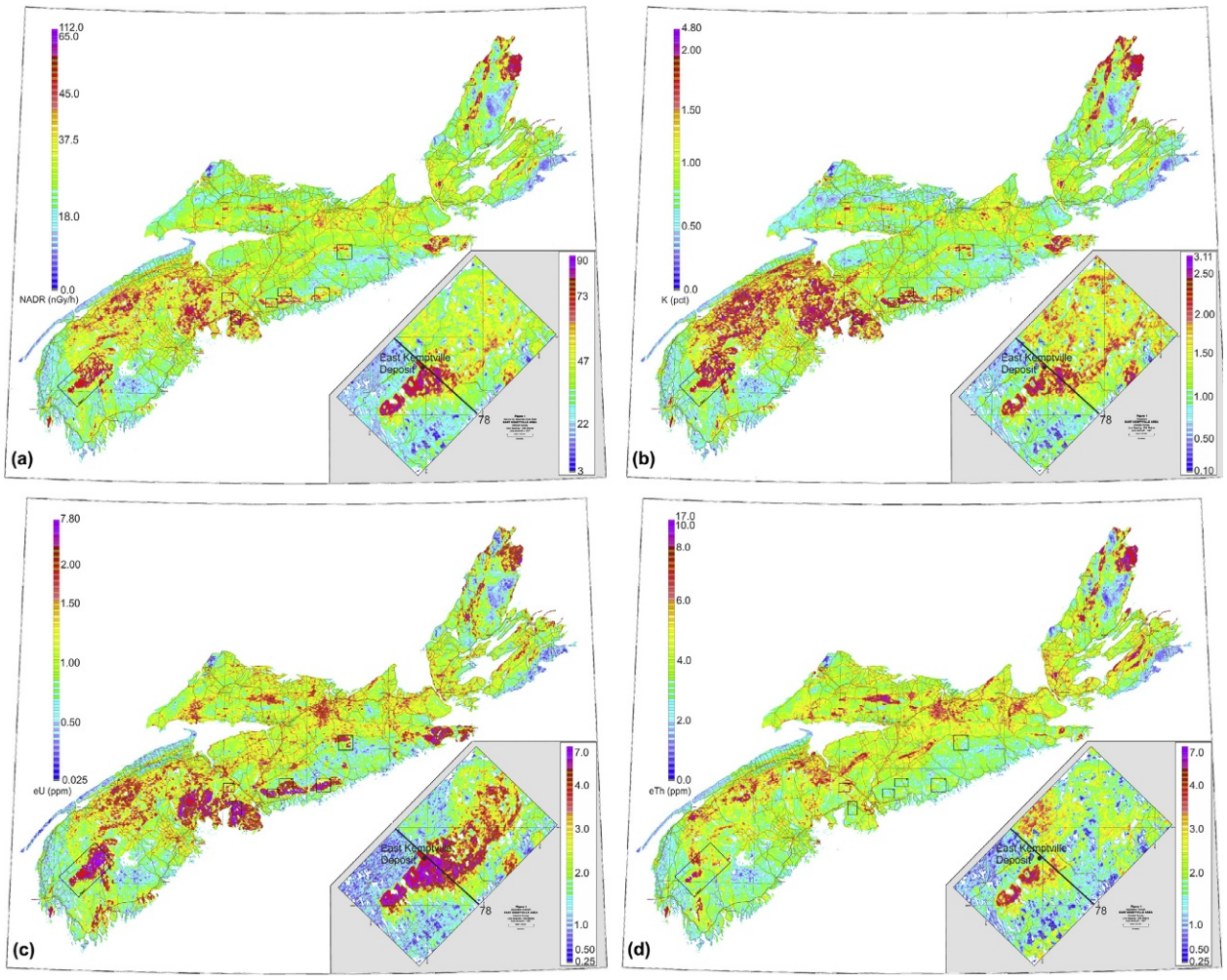


Figure 14 Samples of standard color interval maps and a ternary radioelement image offered by the Geological Survey of Canada for the province of Nova Scotia produced from 1000 m line spaced AGRS data (Carson et al., 2003). Inset maps produced from 250 m line spaced AGRS data for the East Kemptville tin deposit area outlined by the black rectangle in the bottom left corner of the larger maps. (a) Natural air absorbed dose rate (NADR, nGy h^{-1}); (b) potassium (K, %); (c) equivalent uranium (eU, ppm); (d) equivalent thorium (eTh, ppm);

(Continued)

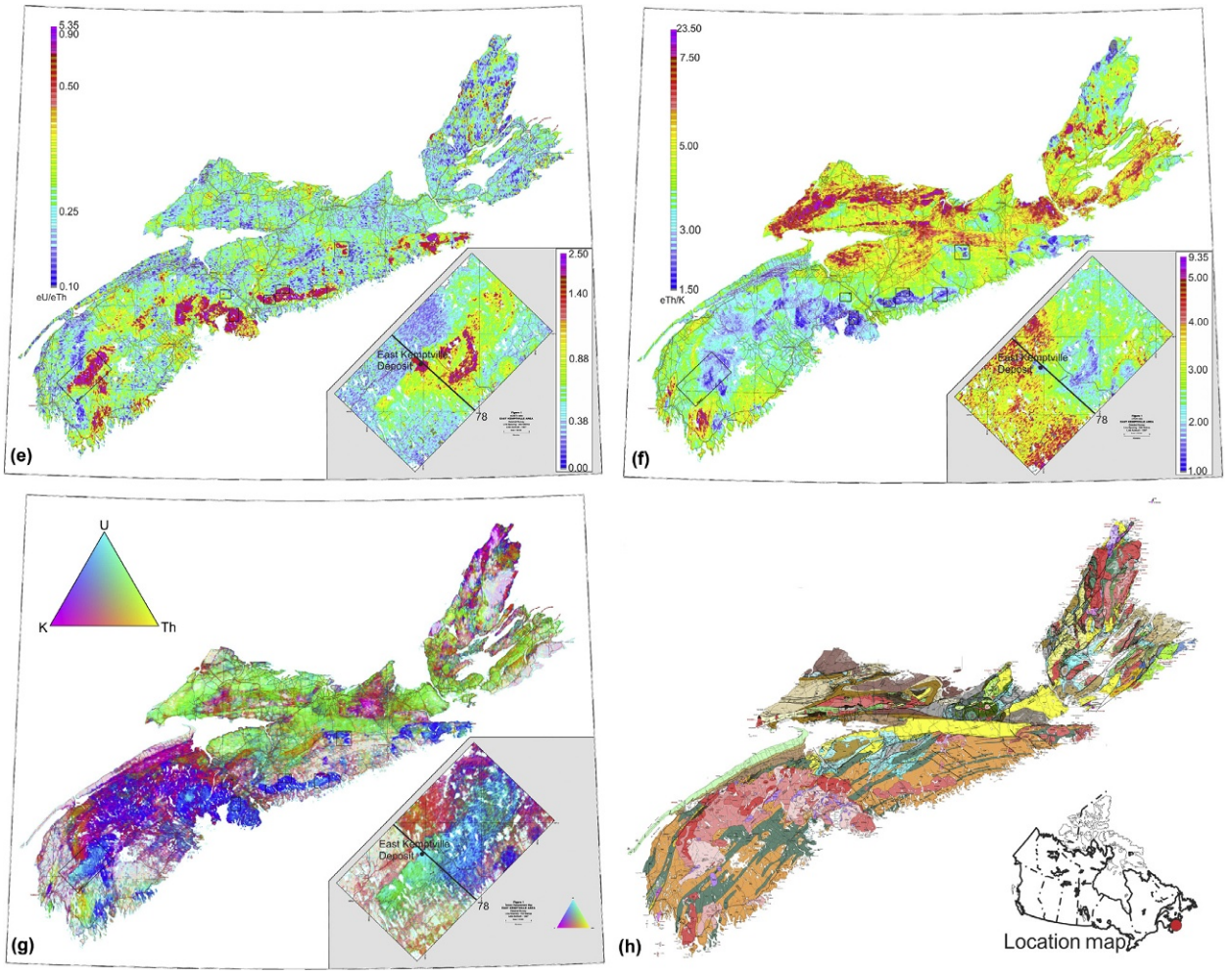


Figure 14 (Continued) (e) equivalent uranium/equivalent thorium ratio (eU/eTh); (f) equivalent thorium/potassium ratio (eTh/K, $\times 10^{-4}$); (g) ternary radioelement image (reproduced (a)-(g) from Carson JM, Holman PB, Ford KL, Grant JA, and Shives RBK (2003) Airborne gamma ray spectrometry compilation, Nova Scotia. Geological Survey of Canada, Open Files 4464-4471, scale 1:500 000); (h) geology map. Reproducd from Keppie JD (compiler) (2000) Geological map of the Province of Nova Scotia. Nova Scotia Department of Natural Resources, Minerals and Energy Branch, Map ME 2000-1, scale 1:500 000.

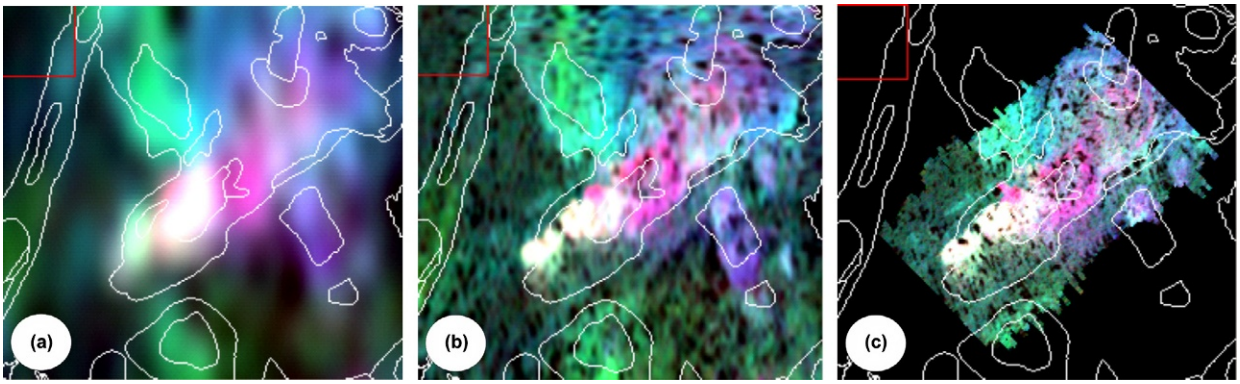


Figure 15 Ternary radioelement images for data acquired with different line spacings over part of the South Mountain Batholith in the area of the East Kemptville tin deposit in southwestern Nova Scotia. (a) 5000 m line spaced data in which eU (ppm), eTh (ppm), and K (%) have been displayed through red, green, and blue colors, respectively; (b) 1000 m line spaced data; (c) 250 m line spaced data. Note the difference in the detail provided by the higher-resolution data. However, the overall patterns of the radioelement signatures are similar, regardless of spatial resolution. Reproduced from Ford KL, Keating P, and Thomas MD (2007) Overview of geophysical signatures associated with Canadian ore deposits. In Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 939–970. Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

Table 6 Correlation between AGRS data derived from different line spacing for eU (ppm), eTh (ppm), and K (%)

	eU – 5 km	*eU – 1 km*	*eU – 0.25 km*
eU – 5 km		0.84	0.76
eU – 1 km			0.87
eU – 0.25 km			
	eTh – 5 km	*eTh – 1 km*	*eTh – 0.25 km*
eTh – 5 km		0.75	0.63
eTh – 1 km			0.85
eTh – 0.25 km			
	%K – 5 km	*%K – 1 km*	*%K – 0.25 km*
%K – 5 km		0.82	0.63
%K – 1 km			0.86
%K – 0.25 km			

Sample data are from southwestern Nova Scotia, the same data as used in **Figure 15**. Source: Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

Schetselaar (2000) used cross plots of AGRS flight line data to determine if radioelement domains defined by interpretation of gridded data were verifiable through analysis of the flight line data.

11.14.6.3 Visual Enhancement of AGRS Data

With the increasing availability of digital datasets and the standard use of GIS and image processing techniques, establishing relationships between AGRS data and other datasets prior to field mapping or mineral exploration activities or as part of an RPM study has become increasingly popular (Harris et al., 1994; Schetselaar et al., 2008).

11.14.6.3.1 Single band

Because AGRS data usually vary over a narrow range of values compared to other geophysical or remotely sensed data, a number of visual enhancement products can be used to assist in interpretation of the data. AGRS enhancements often include some combination of linear, histogram equalization, and interactive contrast stretching of either a single band (K, eU, or eTh) or ratio (eU/eTh, eU/K, or eTh/K) displayed as an image in concentration units or contrast stretched to a byte format usually to improve the dynamic range of the data. A user should be aware that some enhancements may result in loss of the original concentration or ratio values making it difficult to assign an absolute value to a particular location or color. If the goal is to produce an image in which potentially significant geologic or exploration features are enhanced through the use of a wider range of colors, then the loss of absolute values over the enhancement of relative variations will be acceptable if the original concentration values are preserved in other images. These enhanced products can then be displayed as either black and white images or color images (**Figure 18**). The advantage of color images is that the human eye can distinguish a much larger range of colors compared to black and white tones.

11.14.6.3.2 Multiband

As with optical LANDSAT and/or multibeam radar data, AGRS variables (bands) can be displayed as ternary color composite images similar to those in **Figures 13** and **14**, allowing for the interpretation of three channels of data (K, eU, and eTh) or derived ratios. A computer display system using an additive mix of the primary colors (red–green–blue) similar to single band presentations and multiband or ternary enhancements can also be applied to radioelement ratios (**Figure 19**).

11.14.6.3.3 Enhancement

As mentioned in **Section 11.14.6.3.1**, AGRS data are often characterized by a very narrow range of values compared to other geophysical or remotely sensed data. A variety of enhancement techniques that are commonly applied to LANDSAT data to improve the dynamic range can also be applied to AGRS data to amplify potentially significant yet subtle variations. Commonly used techniques include decorrelation and saturation stretches (Gillespie et al., 1986), available with most image analysis systems. These stretches are more than simple arithmetic transforms that use simple mathematical functions (linear, square root, etc.) to transform the data. They involve either a multivariate statistical transform of the data (decorrelation stretch) or a color-space transform (saturation stretch). In both cases, the goal is the same, to provide an enhanced image in which potential geologic features are enhanced though a wider range of hues. **Figure 20** shows an example of a decorrelation stretch applied to the same data shown in **Figures 18** and **19** for the southern part of the Melville Peninsula in Nunavut, Canada.

11.14.6.3.4 Integrated images

Harris et al. (1994) noted that the enhancement of geophysical data (including AGRS data) by integration with other remotely sensed data often yields geologically significant information that is not readily evident in images or maps derived from the individual datasets. Like optical and microwave data, AGRS data may be superimposed or overlain with a variety of other data creating images that offer complementary information from the combined datasets in one image (Ford et al., 2008a,b; IAEA, 2003; Milligan and Gunn, 1997; Schetselaar et al., 2008). Some of the datasets commonly integrated with AGRS data include LANDSAT Thematic Mapper, radar (i.e., RADARSAT and Shuttle Radar Topographic Mission (SRTM)) data, digital elevation model (DEM) data, hyperspectral data, and magnetic data.

A visual integration of gamma ray and magnetic data (**Figure 21(a)**) often provides an image in which geochemical information from the AGRS data combined with structural features from the magnetic data can be used to define geologic and/or tectonic domains, faults, lineaments, granitic rocks, and compositional variations within those granitic rocks and various other lithologic units. AGRS data can also be integrated

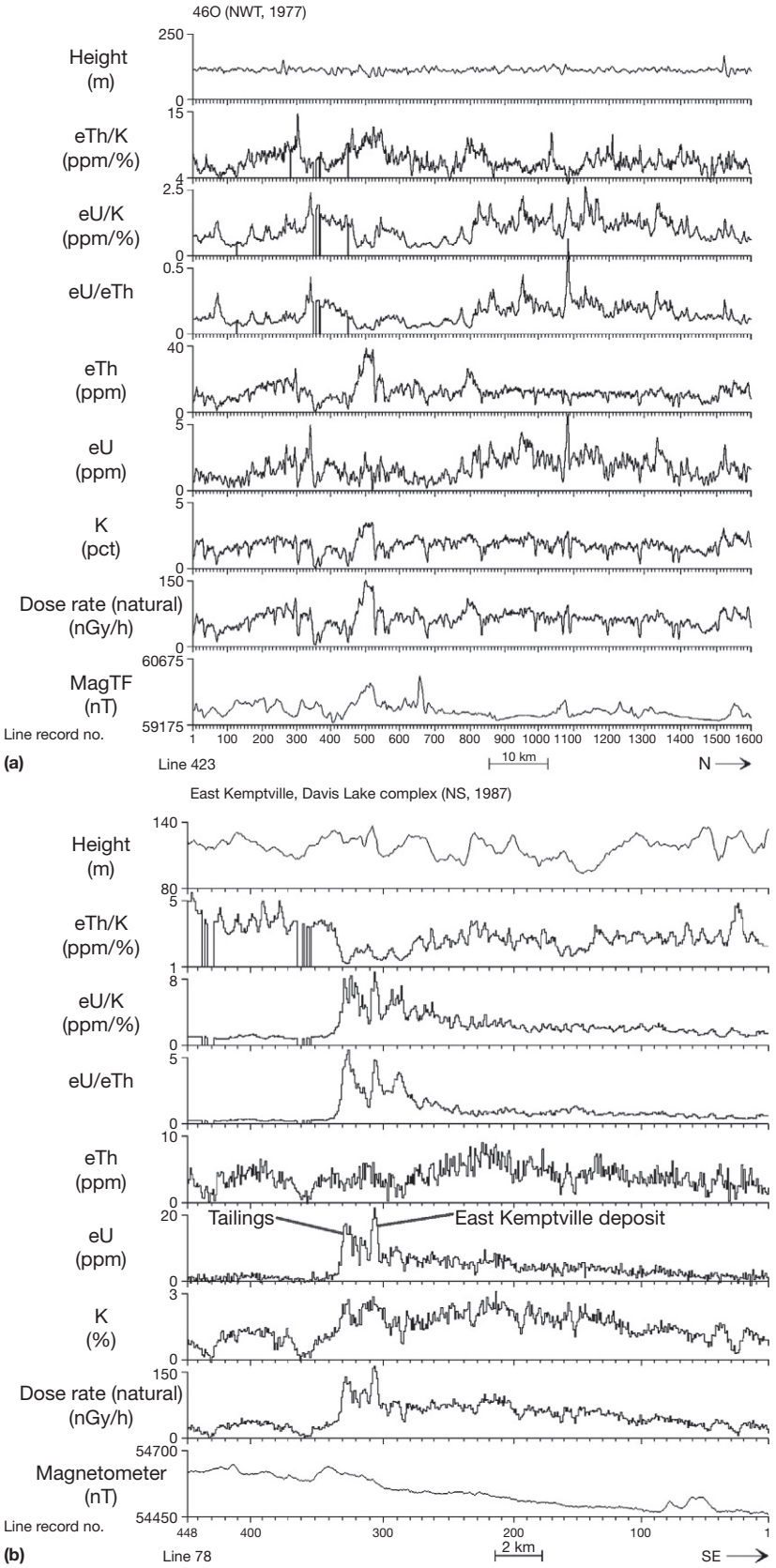


Figure 16 Samples of airborne gamma ray spectrometric multivariable stacked profiles. (a) Line 423 in the southern Melville Peninsula area, Nunavut (see **Figure 13** for location). Note two significant eU and eU/eTh ratio anomalies that can be lost if only the total radioactivity data are presented. (b) Line 78 from over the East Kemptville tin deposit in southwestern Nova Scotia (see **Figure 14** for location). Note the high eU and eU/eTh ratio signature associated with the tin deposit, and the progressive decrease of eU and increase of eTh concentrations in the down-ice dispersal direction (left to right) away from the deposit anomaly. Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643.

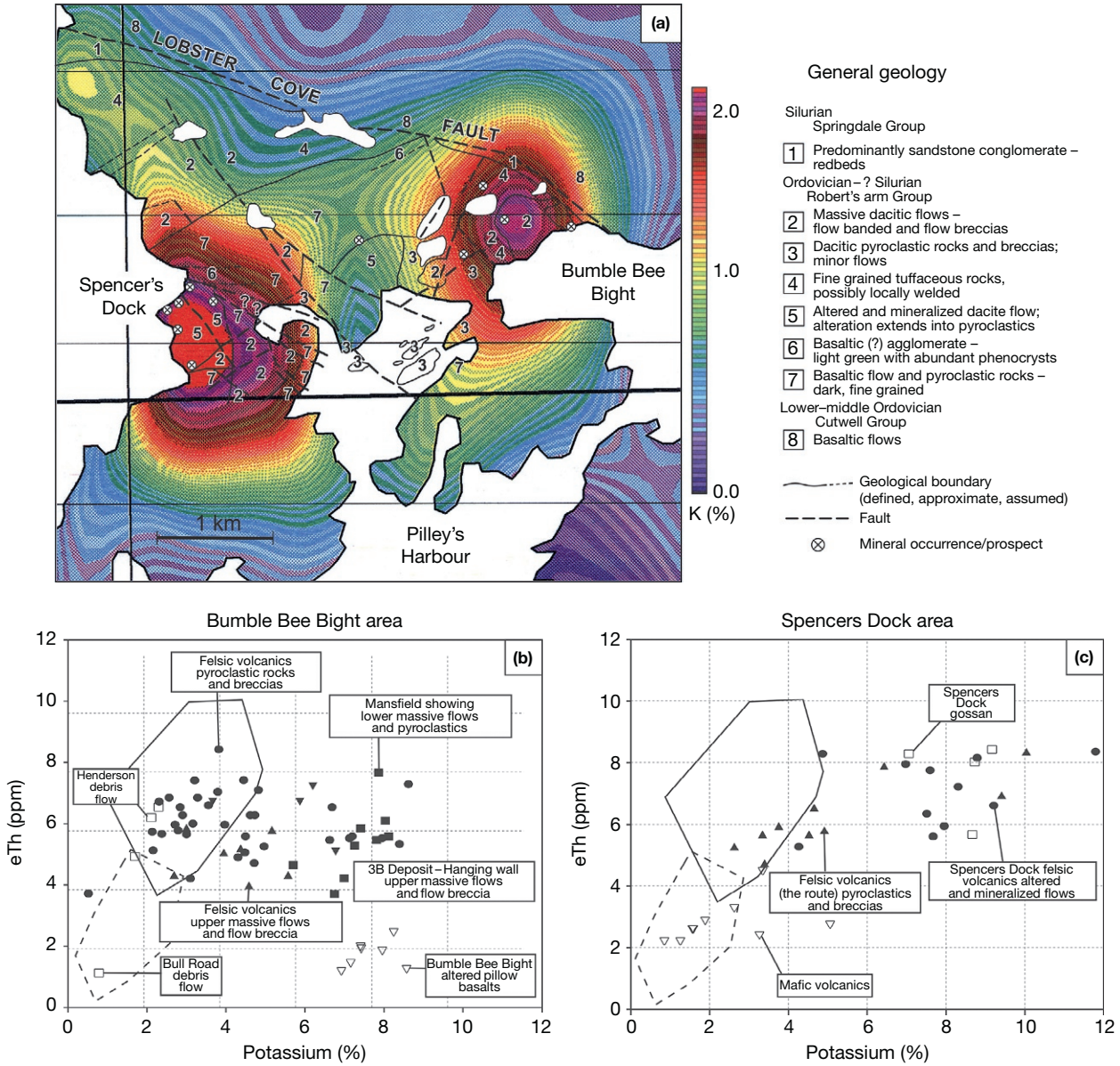

(a)
LOBSTER COVE FAULT
Spencer's Dock
Bumble Bee Bight
Pilley's Harbour
1 km
2.0
1.0
0.0
K (%)
General geology
Silurian
Springdale Group
1 Predominantly sandstone conglomerate – redbeds
Ordovician – ? Silurian
Robert's arm Group
2 Massive dacitic flows – flow banded and flow breccias
3 Dacitic pyroclastic rocks and breccias; minor flows
4 Fine grained tuffaceous rocks, possibly locally welded
5 Altered and mineralized dacite flow; alteration extends into pyroclastics
6 Basaltic (?) agglomerate – light green with abundant phenocrysts
7 Basaltic flow and pyroclastic rocks – dark, fine grained
Lower–middle Ordovician
Cutwell Group
8 Basaltic flows
Geological boundary (defined, approximate, assumed)
Fault
Mineral occurrence/prospect
Bumble Bee Bight area
(b)
Felsic volcanics pyroclastic rocks and breccias
Mansfield showing lower massive flows and pyroclastics
Henderson debris flow
3B Deposit – Hanging wall upper massive flows and flow breccia
Felsic volcanics upper massive flows and flow breccia
Bull Road debris flow
Bumble Bee Bight altered pillow basalts
eTh (ppm)
Potassium (%)
Spencers Dock area
(c)
Spencers Dock gossan
Felsic volcanics (the route) pyroclastics and breccias
Spencers Dock felsic volcanics altered and mineralized flows
Mafic volcanics
eTh (ppm)
Potassium (%)

Figure 17 Cross plots showing the distribution of GGRS measurements symbol-coded by lithology for eTh (ppm) versus K (%) from the Pilley's Island massive sulfide deposit, Newfoundland. Dashed line indicates the range of K and eTh concentrations for equivalent unaltered mafic volcanics. The solid line indicates the range of K and eTh concentrations for equivalent unaltered felsic volcanics. Note the trend toward high K concentrations reflecting K enrichment related to alteration associated with massive sulfide mineralization. Reproduced from Ford KL, Keating P, and Thomas MD (2007) Overview of geophysical signatures associated with Canadian ore deposits. In Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 939–970. Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

with topographic data or DEM data to create an image in which gamma ray patterns can be directly compared to topographic patterns (**Figure 21(b)**). This can be applied to help determine the source of AGRS anomalies. AGRS image data can also be integrated with other gamma ray images; for example, eTh/K on K may potentially illustrate the location of potassium alteration zones associated with some types of mineral deposits.

The integration of AGRS data with LANDSAT or radar imagery has proved particularly useful as an enhancement product because, like **Figure 21**, it allows the geochemical and compositional information from the AGRS data to be interpreted along with structural information from the remotely sensed datasets (Schetselaar, 2000; Wilford et al., 1997). In particular, Schetselaar et al. (2008) found that geochemical

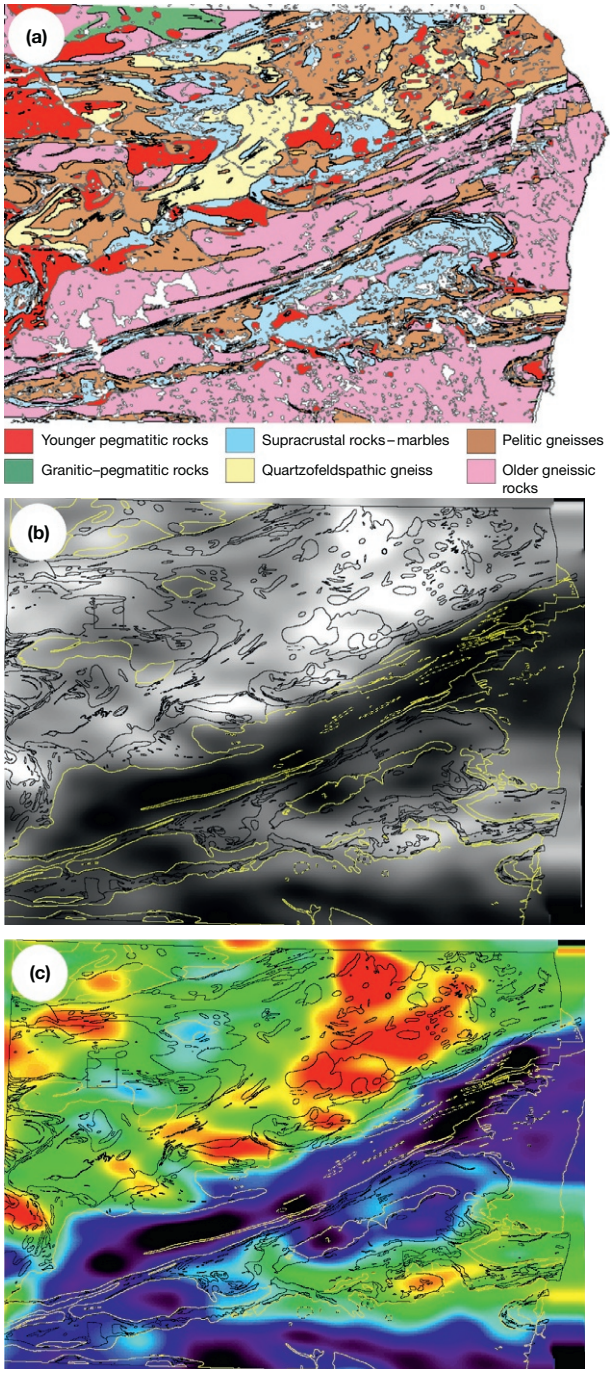


Figure 18 Single band (eU) visualization of AGRS data over part of the southern Melville Peninsula, Nunavut (see **Figure 13** for location). (a) Bedrock geology from Henderson (1987); (b) black and white image of eU data; (c) color image of same eU data. Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

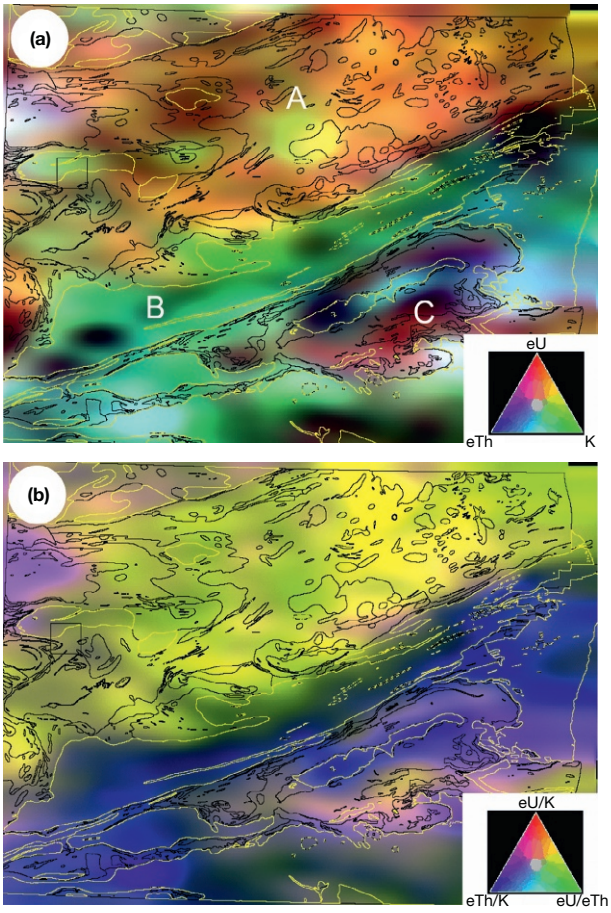


Figure 19 Multiband or color composite (ternary) AGRS images over part of the southern Melville Peninsula, Nunavut (see **Figure 18** for geology). (a) Individual radioelements (eU = red, eTh = blue, K = green); (b) radioelement ratios (eU/K = red, eTh/K = blue, eU/eTh = green). Note the visual identification of the major geologic/tectonic terrains (A, pelitic gneisses and pegmatites; B, older gneisses: C, supracrustal sequences). Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

information from a ternary radioelement map integrated with structural information from LANDSAT images and shaded-relief SRTM data proved to be the most useful integration product to support geologic mapping in Mozambique. **Figure 22** illustrates how the intensity of an AGRS ternary image, where eU = red, eTh = green, and K = blue, has been modified by airborne side-looking synthetic aperture radar data to provide a pseudo three-dimensional (3-D) view of the ground surface. This integrated product enhances the division of geologic features, such as the Coldwell Complex into multiple phases based on not only the unique AGRS signature but also the unique textural signature (Graham et al., 1994).

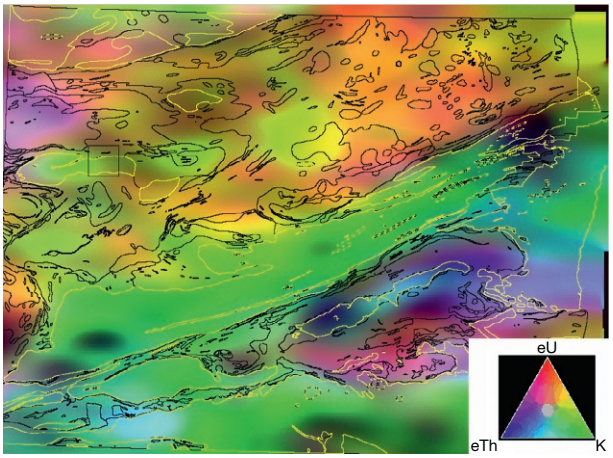


Figure 20 Color composite (ternary) AGRS image over part of the southern Melville Peninsula, Nunavut (see **Figure 18** for geology), showing results of a decorrelation stretch that takes the elemental (K, eU, and eTh) data, transforms the data using the principal component analysis (PCA) transform, stretches the histogram of each component image, and then applies an inverse PCA transform recreating stretched versions of the data, which are then displayed through colors red–green–blue. Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008a) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

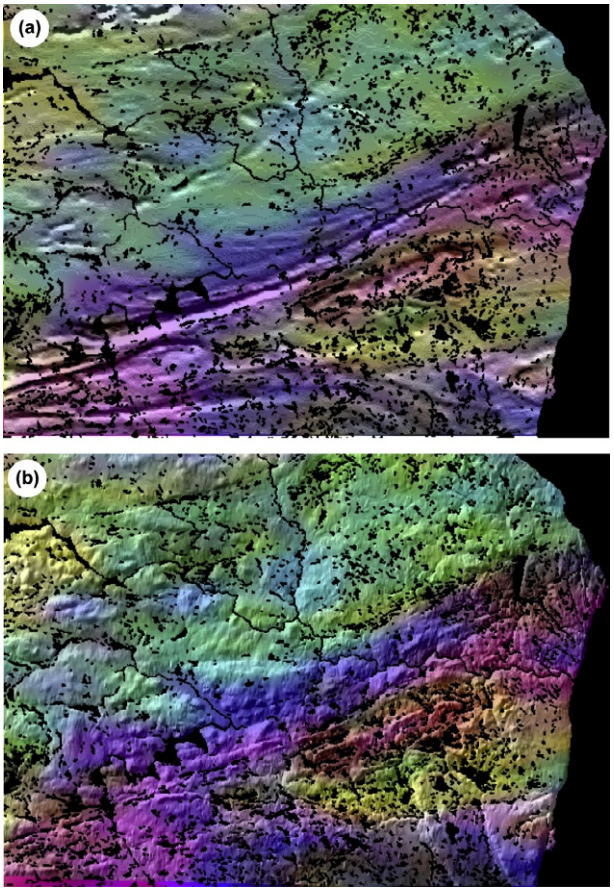


Figure 21 Examples of integrated (fused) visualizations of magnetic and topographic (digital elevation model, DEM) data with AGRS data from the southern Melville Peninsula, Nunavut (see **Figure 18** for geology). (a) AGRS ternary image fused with shaded total field magnetic data showing geochemical information from the AGRS data combined with structural features from the magnetic data, (b) AGRS ternary image fused with DEM showing topographic patterns that can be directly related to radioelement signatures. Adapted from Schetselaar EM, Harris JR, Lynds T, and de Kemp EA (2007) Remote predictive mapping 1. Remote predictive mapping (RPM): A strategy for geological mapping of Canada's North. *Geoscience Canada* 34: 93–111.

11.14.6.3.5 Dimensional visualizations

Different perspective views, for example, 2.5-D or 3-D, can be employed to not only fuse different data types but also effectively visualize the relationships (both spatial and signature response) between them (Ford et al., 2008a,b; Wilford et al., 1997). **Figure 23** shows a 2.5-D perspective image in which the relief reflects the magnetic total field and the color reflects the eU response. Correlations between the two data types can be seen over supracrustal rocks labeled (a) and (b) in **Figure 23**.

11.14.6.4 Data Presentation and Visualization of BGRS Data

Data presentation of borehole gamma ray or gamma ray spectral logs generally consists of plotting the TC gamma ray activity or radioelement content versus depth along the borehole. The standard practice is to plot the depth axis vertically and the TC, potassium, uranium, and thorium values on the horizontal axis.

A TC gamma ray logging system can only display a single log, that is, the TC log. For a spectral gamma ray logging system, generally, four logs are presented: the TC log and the three radioelement logs (K, eU, and eTh). Sometimes, the ratios of the radioelements, eU/eTh, eU/K, and eTh/K, are presented. If logs are available from several drill holes along a section, then any one of the logging parameters, TC or the radioelements, may be contoured from hole to hole. Geologic or lithologic information may be superimposed if available and geologic columns may also be presented. A TC gamma ray log from a kimberlite pipe is shown in **Figure 24**.

11.14.7 Data Analysis

11.14.7.1 Introduction

Analysis of GRS data depends on the application and datasets available, including not only radiometric (total radioactivity versus spectral) but also other geophysical data, for example, magnetic data, and geologic or remotely sensed data such as radar or hyperspectral data. Data analysis methods for AGRS are discussed first followed by BGRS data analysis methods. There are a number of data analysis techniques that can be

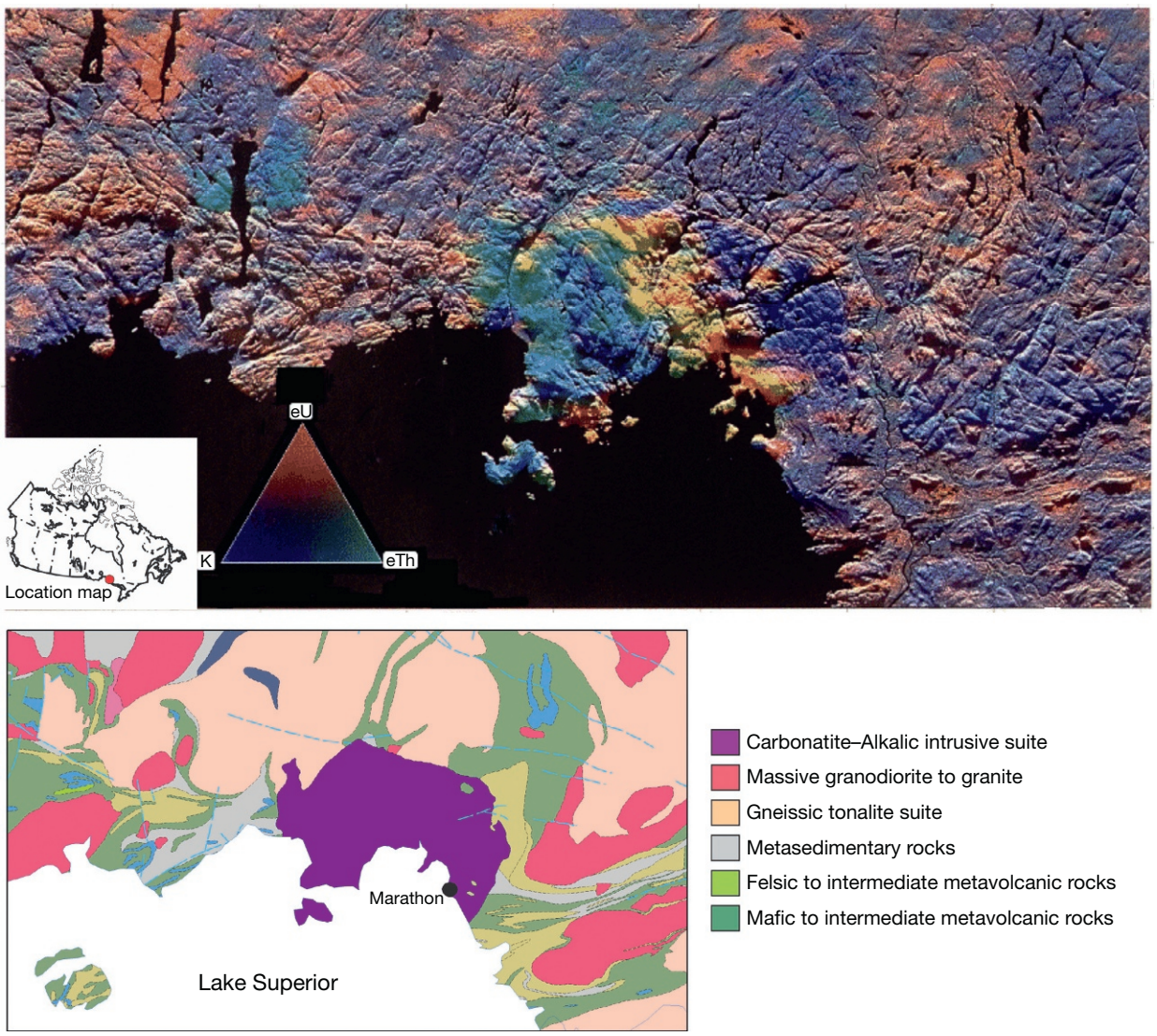


Figure 22 Example of an integrated (fused) visualization product of AGRS elemental data (K, eU, and eTh) displayed through colors blue–red–green, respectively, combined with wide swath radar data showing division of the alkalic Coldwell Complex into unique radioelement and textural domains. Geologic inset adapted from Ontario Geological Survey (1991) Bedrock geology of Ontario, West Central Sheet; Ontario Geological Survey, Map 2542, scale 1:1 000 000.

applied to AGRS data (IAEA, 2003; Schetselaar et al., 2007). All of these methods can be used to augment visual interpretation of AGRS data with the goal of improving interpretation results. IAEA (2003) lists a number of these data analysis methods including mean differencing, regression analysis, and PC analysis (PCA). As mentioned earlier, AGRS data comprise seven channels of data, including ratios; thus, multivariate processing techniques can be applied to the data for both enhancement and classification purposes. Ford et al. (2008a,b) demonstrated how the PCA transform can be applied to AGRS data to make a reduced set of data comprising a series of component images that are a linear combination of all the AGRS data. Individually, these component images often reveal geologic information (**Figure 25**), but when combined into a color three-component or ternary image (**Figure 26**), geologic features may be further enhanced by variations in color.

11.14.7.2 Classification

Data analysis techniques can be used to classify AGRS data into radioelement domains defined as areas with similar radioelement signatures or associations. The same classification techniques that are applied to LANDSAT and radar data can also be applied to AGRS data to create predictive maps. Classification of AGRS data into radioelement domains can be done manually based on a visual interpretation (Campbell et al., 2007), semi-automatically or automatically. Schetselaar et al. (2007) stated that visual interpretation methods are often judged to outperform automated domain recognition methods. However, they also state that automated methods are superior for manipulating and interpreting multiple datasets that have a large number of variables including AGRS data. **Figure 27** shows results of the visual interpretation approach (Schetselaar, 2000; Schetselaar

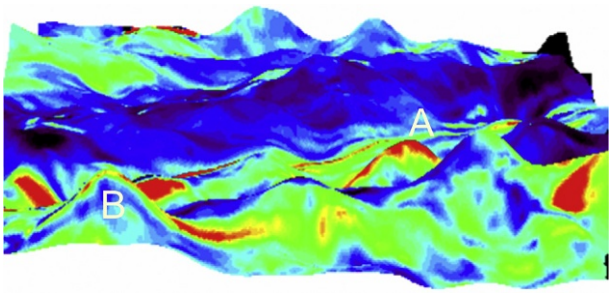


Figure 23 Example of an integrated (fused) 2.5-D visualization of total field magnetic data that provides 'topographic relief' with AGRS eU data providing color variations. Features labeled 'A' and 'B' indicate areas of correlation between high magnetic and high eU signatures. Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

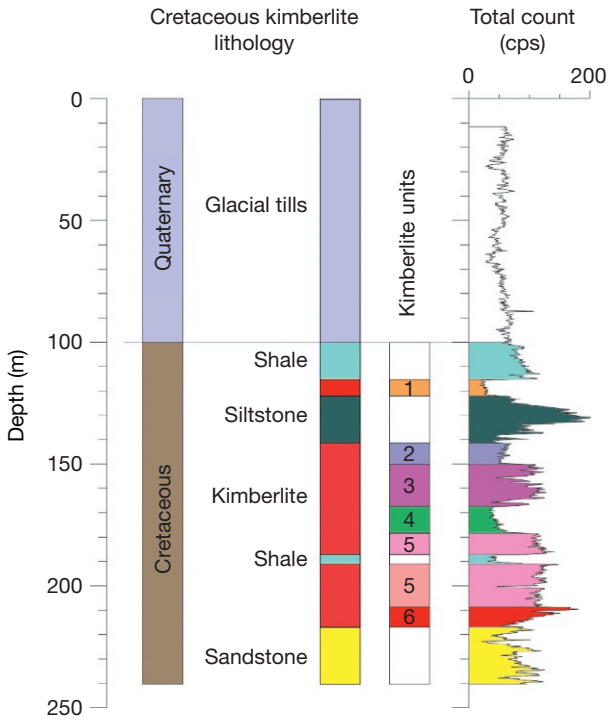


Figure 24 Geologic columns and total count gamma ray log with superimposed interpreted geologic information from a kimberlite pipe in Fort à la Corne cluster, Saskatchewan. Reproduced from Mwenifumbo CJ and Kjarsgaard BA (2000) Gamma ray logging and radioelement distribution in the Fort à la Corne pipe 169. *Exploration and Mining Geology*, *Journal of the Geological Society of the Canadian Institute of Mining* 8: 133–147.

et al., 2007) to produce a remote predictive map to support bedrock mapping. In this example, a remote predictive map was produced from the visual interpretation of both AGRS and aeromagnetic data for the Taltson Magmatic Zone, northeast Alberta. Schetselaar (2000) found that the AGRS data were particularly effective for separating the various granitic rocks and internal compositional variations within them. Interpretation of the magnetic data permitted separation of the granitic rocks from older, gneissic basement units and enclaves of supracrustal rocks within the gneissic basement.

11.14.7.2.1 Unsupervised classification

Harris (1989) stated that a goal of any clustering technique is to classify complex multivariate data into a smaller number of tractable units and produce a predictive map that will reveal patterns that can be directly related to lithologic variations. Various combinations of the AGRS data channels can be classified using an unsupervised classification algorithm to produce a thematic map of radioelement domains or classes that show distinct statistical separation. The advantage of this approach is that it requires little input by the geologist other than specifying the number of classes.

Harris (1989) found that using a migrating means unsupervised clustering algorithm was an effective technique to classify regional AGRS data acquired with 1000 m line spacing, into similar and spatially continuous groups or domains. The number of clusters and the initial mean value for each cluster for each AGRS channel may be specified or automatically selected by the computer. The use of some a priori knowledge of the AGRS data to provide initial mean values is often desirable as it may result in more meaningful classes. In this example, the radioelement domains produced by clustering AGRS data showed a fairly close spatial correlation with mapped geology and identified other clusters that reflected previously unrecognized compositional variations that elsewhere, in the same geologic terrain, were found to have exploration significance (Ford, 1993). The modifying effects of glacial dispersion could also be quantified by the clustering procedure.

Unsupervised classification using cluster algorithms is often used when there are no field observations, such as GGRS, till geochemistry, and other reliable geologic information. An example of unsupervised classification using reconnaissance AGRS data acquired with 5000 m line spacing is shown in **Figure 28** (Ford et al., 2008a,b; Schetselaar et al., 2007). This example, from the area indicated in **Figure 13**, shows two radioelement domain maps derived from K, eU, and eTh data and their ratios along with the corresponding mean and standard deviation values for each domain. The divisions between the different lithologic/tectonic regimes can be clearly seen on these unsupervised maps.

11.14.7.2.2 Supervised classification

Supervised classification involves the use of training area data that are considered representative of each rock type or surficial unit to be classified. Ford et al. (2008a,b) presented results of a supervised classification (maximum likelihood) applied to reconnaissance (acquired with 5000 m line spacing) AGRS data (**Figure 29**). Maximum likelihood is one of several commonly used algorithms where input for classes established from training site data is used to calculate appropriate statistics (mean and variance–covariance) and a probability function. Each AGRS data pixel is then analyzed to determine the class into which it most likely belongs. In this example, a variety of combinations of AGRS variables including ratios were used to

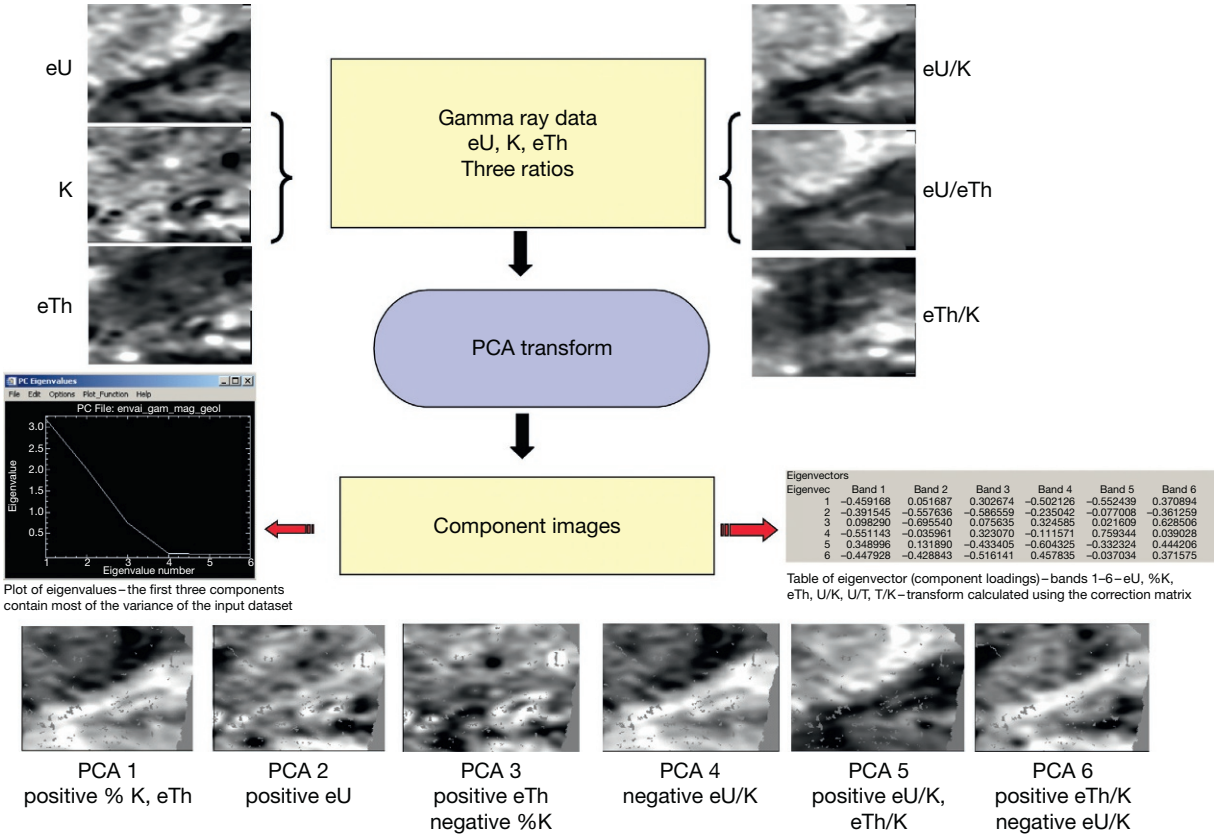


Figure 25 A description of the PCA transform applied to AGRS data over the southern Melville Peninsula, Nunavut (see **Figure 13** for location and **Figure 18** for geology). Elemental (K, eU, and eTh) and ratio (eU/K, eU/eTh, and eTh/K) data are transformed using PCA creating component images (PCA1–PCA6), which are linear combinations of the input data. The first few components contain most of the information (variance) of the dataset (see table of eigenvalues). Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

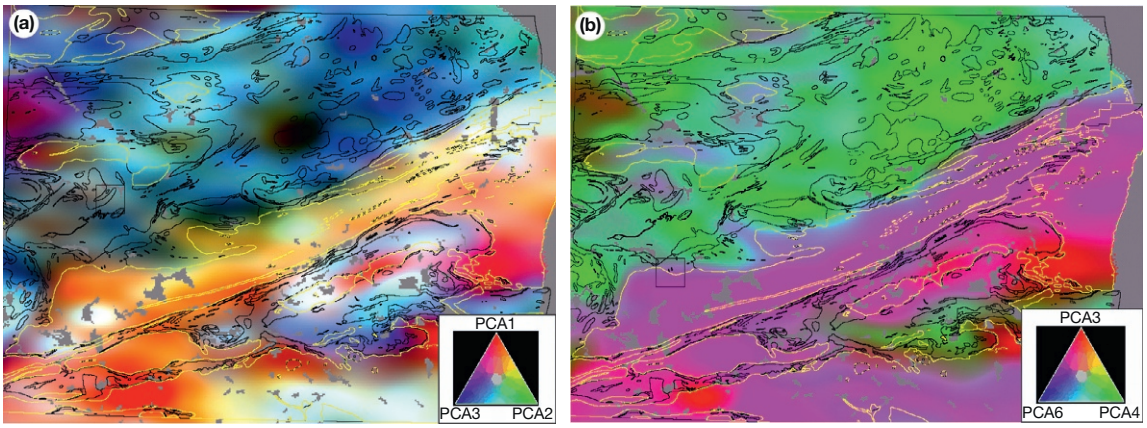


Figure 26 Ternary images of the PCA transform for AGRS data over the southern Melville Peninsula, Nunavut (see **Figure 13** for location and **Figure 18** for geology). (a) PCA 1, 2, and 3 (red, green, and blue); (b) PCA 3, 4, and 6 (red, green, and blue). These ternary images comprising the PCA component images provide enhancement of major geologic and tectonic units. Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

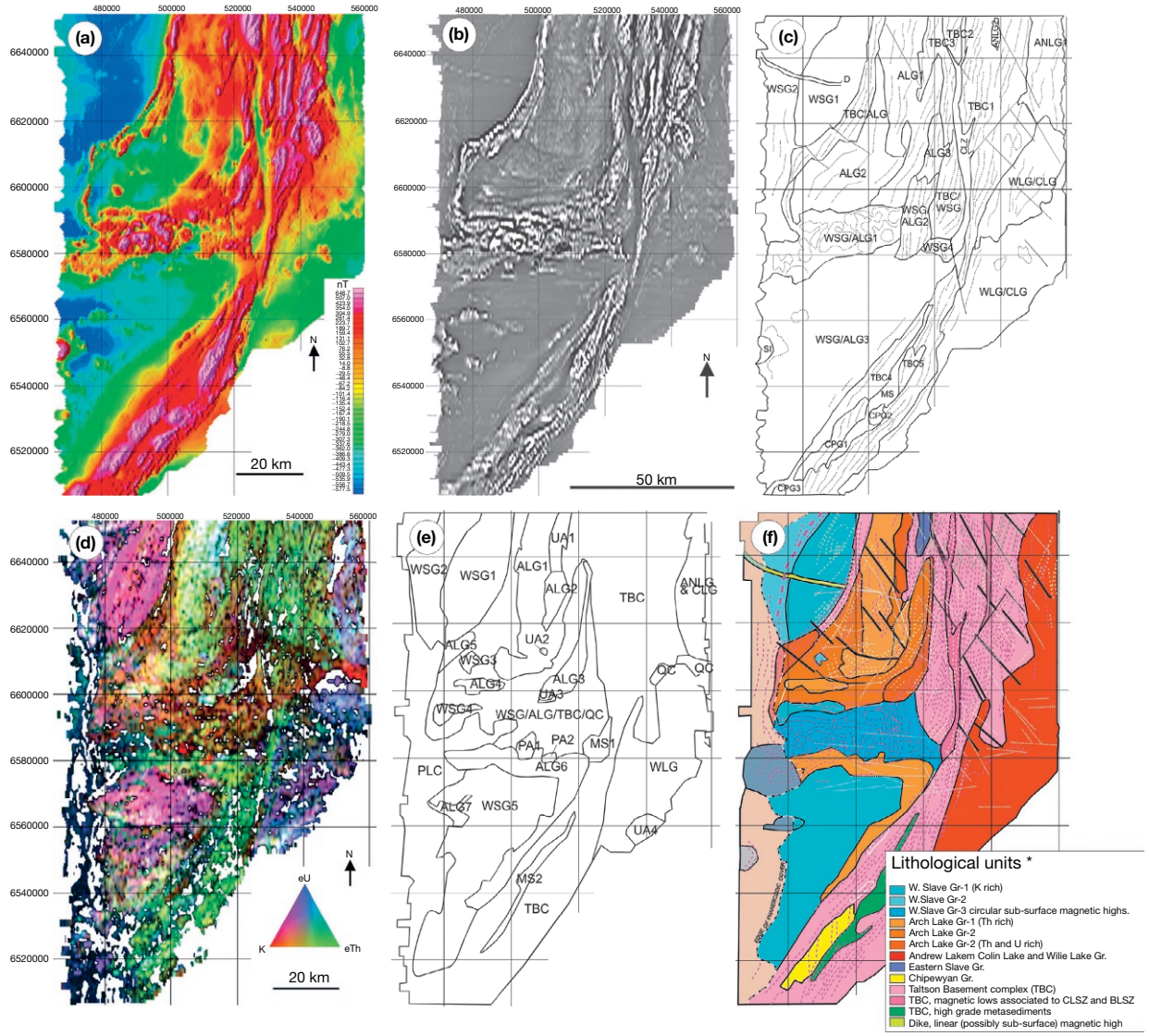
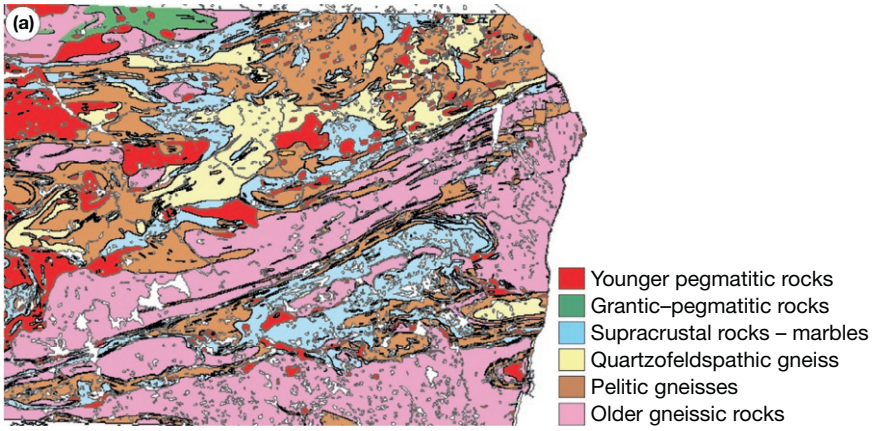


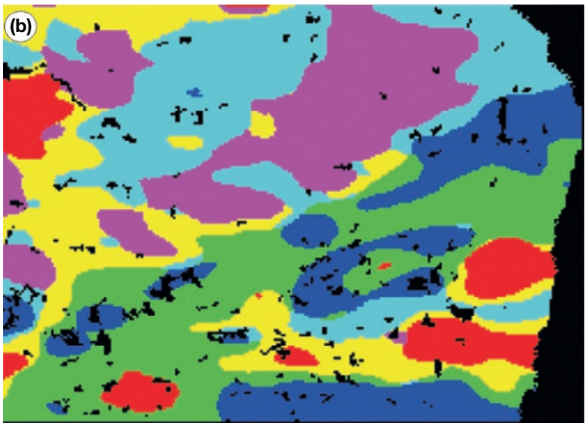
Figure 27 Example of a visual interpretation approach to producing a remote predictive map of the Taltson Magmatic Zone, northeast Alberta; (a) color shaded relief of the total field magnetic data; (b) gray scale image of first vertical derivative magnetic data; (c) interpretation of total field and first vertical derivative magnetic data; (d) ternary radioelement image of AGRS data (K = red, eTh = green, and eU = blue); (e) interpretation of radioelement domains; (f) remote predictive map based on combined interpretation of magnetic, AGRS, LANDSAT (not shown), and ERS-1 (not shown) data. Map grid is 20 km by 20 km. Reproduced from Schetselaar EM, Harris JR, Lynds T, and de Kemp EA (2007) Remote predictive mapping 1. Remote predictive mapping (RPM): A strategy for Geological Mapping of Canada's North. *Geoscience Canada* 34: 93–111.

classify the major tectonic domains over the same area shown in Figure 28. Using all the AGRS variables and total field magnetic data resulted in a classification accuracy of approximately 70%. Discrepancies or misclassified areas serve as potential targets for subsequent mapping or mineral exploration activities. Similar classification accuracies (65–70%) were achieved when a variety of supervised classification methods (maximum likelihood, predictive neural network, and artificial neural network) were applied to regional 1000 m line spacing AGRS and magnetic data to map lithologic units in a geologically complex area of northeast Alberta, Canada (Schetselaar, 2000; Schetselaar et al., 2000).

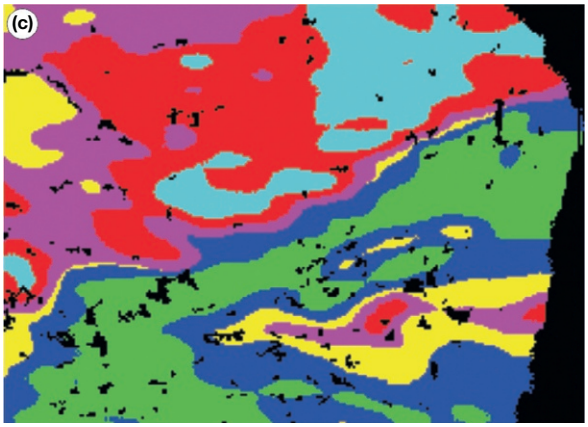
Different supervised classification algorithms are available. Figure 30 presents results of a neural network classification of lithologic units over the same area shown in Figure 28. Despite using reconnaissance AGRS data, classification accuracies of 64% using only the K, eU, and eTh data were achieved. This increased to 70% when total field data were included.

In another example of supervised classification, this time applied to high-resolution (acquired with 500 m line spacing) AGRS data to support geologic mapping in a mountainous, sedimentary-dominated terrain, Harris et al. (2008) showed that classification accuracies of lithologic units in excess of 80% can be achieved using only the AGRS data. This improves to >90% when magnetic and LANDSAT data are included. These results support the use of a variety of geoscience data, each responsive to different lithologic characteristics to produce more accurate classification results.

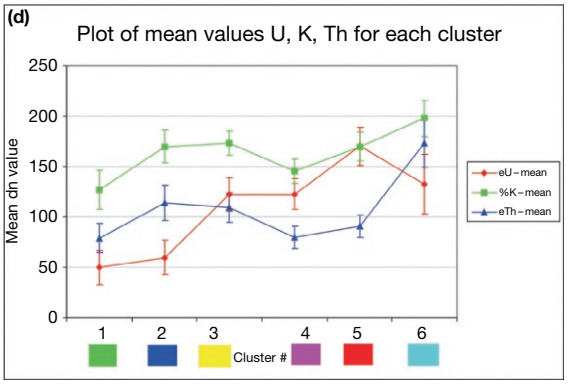




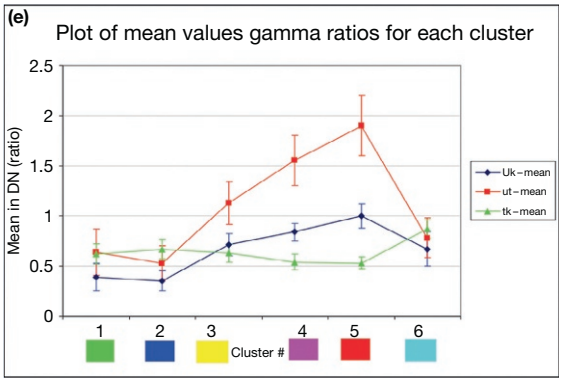
Radioelement map derived from unsupervised classification of the elemental gamma ray data (eU, eTh, K) showing automatic generation of radioelement domains



Radioelement map derived from unsupervised classification using gamma ray spectrometry ratios (eU/eTh, eU/K, eTh/K) showing the automatic generation of radioelement domains



Plot of mean value and standard deviation for each cluster (class) for AGRS elemental data



Plot of mean value and standard deviation for each cluster (class) for AGRS ratio data

Figure 28 Unsupervised classification of AGRS data over the southern Melville Peninsula, Nunavut, showing the automatically generated radioelement domains or classes. (a) Bedrock geology from Henderson (1987); (b) and (c) classification (prediction) maps of radioelement domains derived from elemental (K, eU, and eTh) and ratio (eU/eTh, eU/K, and eTh/K) data, respectively; (d) and (e) means and standard deviations of AGRS responses for elemental and ratio data, respectively, for each radioelement domain. The classification maps display good correspondence with bedrock geology. Areas with unexpected response for a given rock type warrant field checking. Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

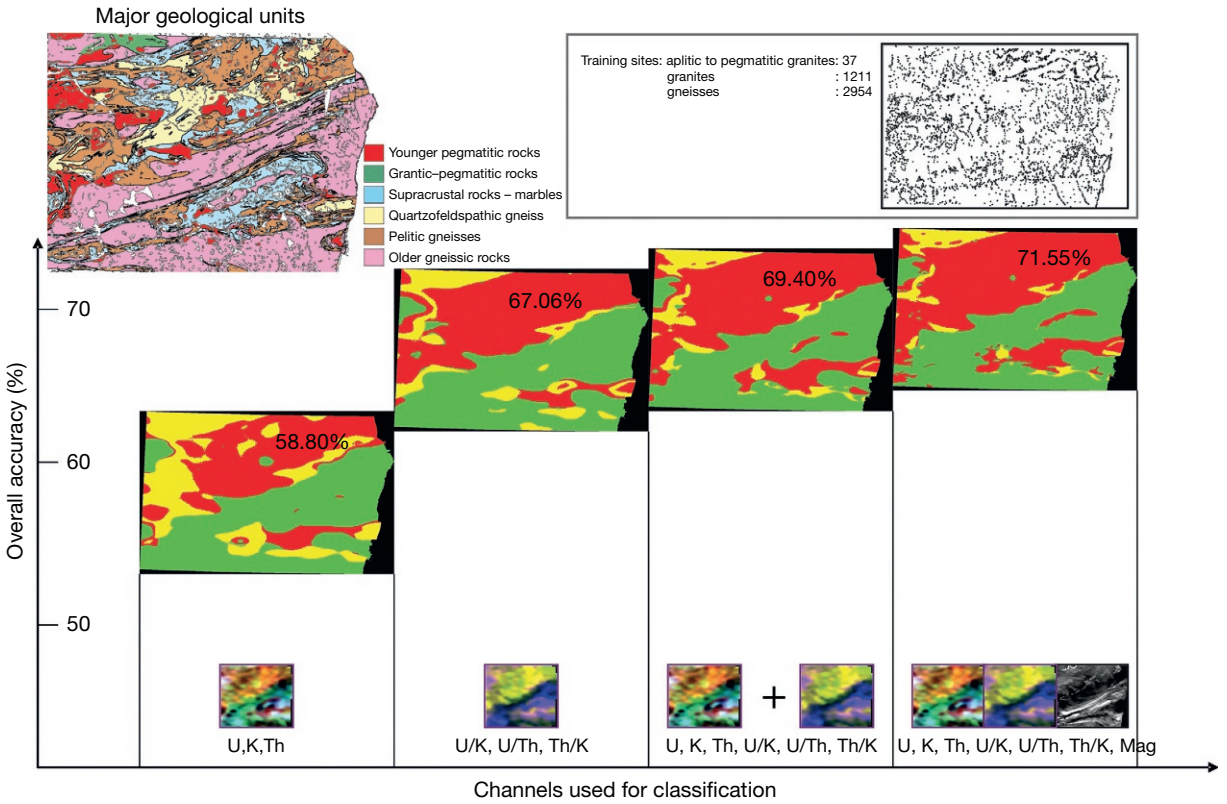


Figure 29 Supervised classification (prediction) of major tectonic divisions using a maximum likelihood algorithm applied to AGRS and total field magnetic data over the southern Melville Peninsula, Nunavut. Training samples for each of the major tectonic divisions were gathered from available field data (point locations). Some were used for classification purposes, while others were used for validation (cross correlation) purposes. Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

11.14.7.3 Analysis of BGRS Data

The approach to data analysis and interpretation of gamma ray logs depends on the application and the available datasets (TC or spectral data). As will be discussed in the following section, the main applications of gamma ray logging data are in geologic mapping (sedimentary, igneous, and metamorphic geologic environments) and in mineral exploration, especially for uranium. In the analysis of gamma ray spectral logs, one of the methods used to enhance the information extracted from the spectral data is to compute ratios of the radioelements. Histograms and cross plots between the three radioelements and their ratios can also be generated.

11.14.7.3.1 TC and spectral log data analysis

Analysis and presentation of TC, K, U, and Th data involve plotting them as log traces, constructing frequency distributions for univariate data, and cross plotting pairs of parameters (e.g., Th and U) in order to determine their relationships and to establish patterns for geologic interpretation (Mwenifumbo, 1993). **Figure 31** is an example showing how cross plots or scatter plots of two radioelements, K and Th, clearly identify trends and clustering associated with Cretaceous sediments and different kimberlite facies in a kimberlite pipe in Saskatchewan (Mwenifumbo and Kjarsgaard, 2000). These observed modes in the distribution can be used to classify unknown lithology in subsequent drill holes.

11.14.7.3.2 TC and spectral log data interpretation

In geologic mapping, log data interpretation primarily consists of integrating the geology with the radiometric data. This is generally accomplished visually, establishing characteristic log responses of the different lithologic/geologic units and noting the relative radioelement abundances (K, U, and Th) and their ratios. These characteristic signatures are then used to establish unit boundaries, for stratigraphic correlation between holes and to establish the presence of geologic features such as faults, unconformities, and depositional environments (diagenesis and alteration) in sedimentary rocks (Serra et al., 1980). Histograms and cross plots of the radioelements are generated to examine the distribution and/or clustering of the data that characterizes the different lithologies. The shape of TC gamma ray logs is often used to establish depositional environments, for example, sediments fining upward or downward. Also, the enrichment of uranium is often an indication of sediments deposited in an acidic, organic-rich environment (Fertl and Rieke, 1979).

In mineral exploration, the TC and spectral gamma ray logs (K, U, and Th) and their ratios provide characteristic signatures useful for mapping prospective lithology and alteration

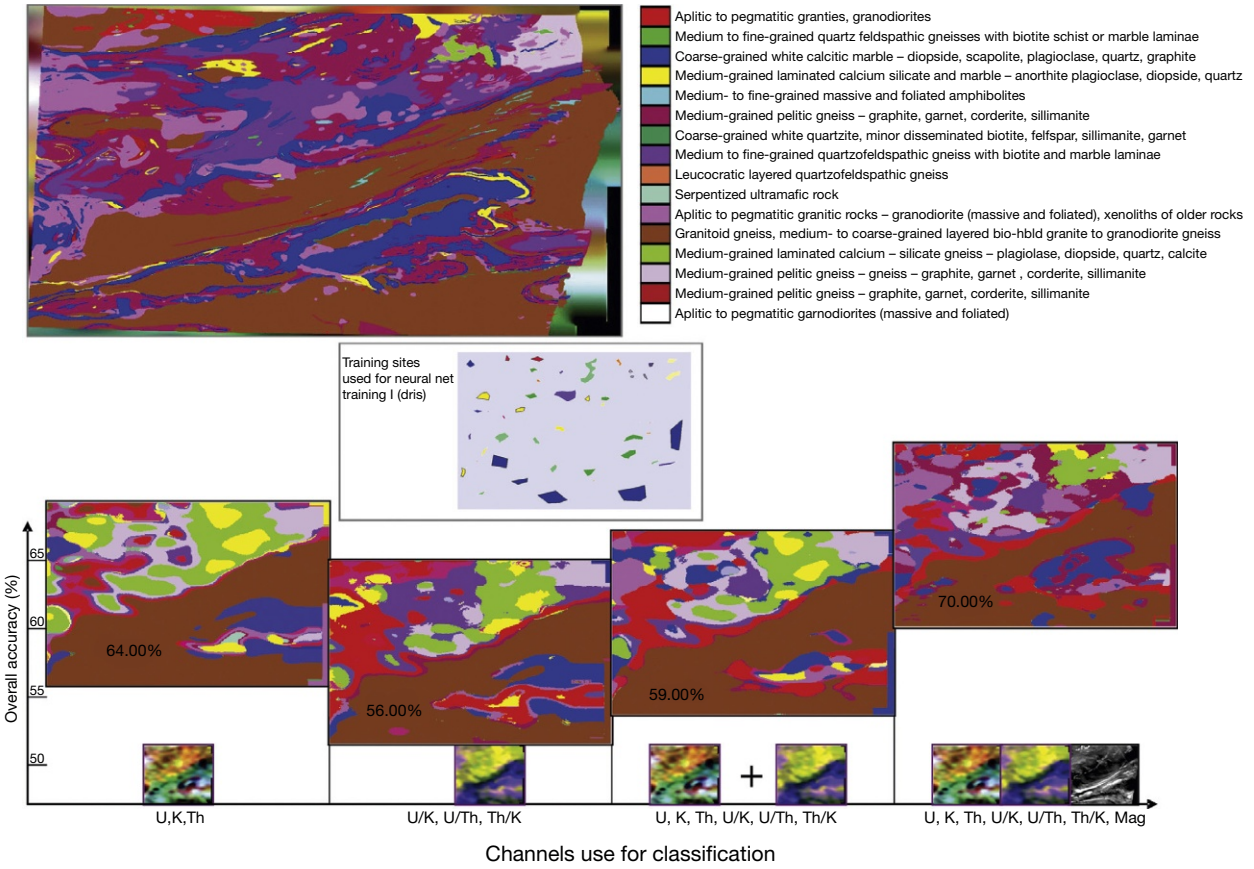

Aplitic to pegmatitic granties, granodiorites
Medium to fine-grained quartz feldspathic gneisses with biotite schist or marble laminae
Coarse-grained white calcitic marble – diopside, scapolite, plagioclase, quartz, graphite
Medium-grained laminated calcium silicate and marble – anorthite plagioclase, diopside, quartz
Medium- to fine-grained massive and foliated amphibolites
Medium-grained pelitic gneiss – graphite, garnet, corderite, sillimanite
Coarse-grained white quartzite, minor disseminated biotite, felfspar, sillimanite, garnet
Medium to fine-grained quartzofeldspathic gneiss with biotite and marble laminae
Leucocratic layered quartzofeldspathic gneiss
Serpentized ultramafic rock
Aplitic to pegmatitic granitic rocks – granodiorite (massive and foliated), xenoliths of older rocks
Granitoid gneiss, medium- to coarse-grained layered bio-hbld granite to granodiorite gneiss
Medium-grained laminated calcium – silicate gneiss – plagiolase, diopside, quartz, calcite
Medium-grained pelitic gneiss – gneiss – graphite, garnet , corderite, sillimanite
Medium-grained pelitic gneiss – graphite, garnet, corderite, sillimanite
Aplitic to pegmatitic garnodiorites (massive and foliated)
Training sites used for neural net training I (dris)
Overall accuracy (%)
65
60
55
50
64.00%
56.00%
59.00%
70.00%
U,K,Th
U/K, U/Th, Th/K
U, K, Th, U/K, U/Th, Th/K
U, K, Th, U/K, U/Th, Th/K, Mag
Channels use for classification

Figure 30 Supervised classification (prediction) of geologic units using a neural network algorithm applied to AGRS and total field magnetic data over the southern Melville Peninsula, Nunavut. Training samples for each of the major units were gathered from the geologic map areas. Some of these were used for classification, whereas others were used for validation or cross correlation. Adapted from Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, ch. 4, pp. 53–97. Ottawa, ON: Geological Survey of Canada, Open File 5643; Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's north. *Geoscience Canada* 35: 109–126.

(e.g., potassic alteration) associated with mineralization. The interpretation process again consists of visual identification of changes in radioelement content and log response characteristics and correlating them with the interpreted lithology or geology. In uranium exploration, the primary objective of gamma ray logging is to locate uranium mineralization (depth and thickness of the mineralization) and to determine uranium ore grade (Mwenifumbo et al., 2004). Quantitative determination of the uranium grade and thickness of uraniferous zones can be obtained by analyzing the count rate anomalies on the gamma ray log as described in **Section 11.14.5**. Logs are therefore presented as the standard depth–gamma count rate profiles along the borehole axis to analyze the anomalies.

More advanced analysis of gamma ray logs includes statistical analysis such as clustering, classification, and change point (Killeen et al., 1997). They are used to guide the geologic interpretation, accurately locate boundaries of geologic units in boreholes, and detect subtle variations in signature within units that geologically appear the same. Various methods of cross plotting of the radioelements (K, U, and Th) have been used to differentiate rocks formed by different geologic processes (e.g., kimberlite pipes and ultramafics), or magmatic differentiation in igneous rocks (basic, mafic, and ultramafic), or sedimentary rock types (Serra et al., 1980).

11.14.8 Interpretation and Applications

11.14.8.1 Introduction

The following section examines a number of important general interpretation issues followed by examples from AGRS surveys that illustrate applications to bedrock and surficial geologic mapping and exploration for a variety of mineral deposit types. This is then followed by a number of examples that illustrate the application of BGRS to stratigraphic mapping and direct and indirect mineral exploration. Finally, some examples of the application of AGRS survey techniques for environmental radiation monitoring are presented.

11.14.8.2 General Comments on Interpretation

All rocks and soils are naturally radioactive, containing varying proportions of a range of radioactive elements. AGRS, GGRS, and BGRS provide methods of measuring concentrations of

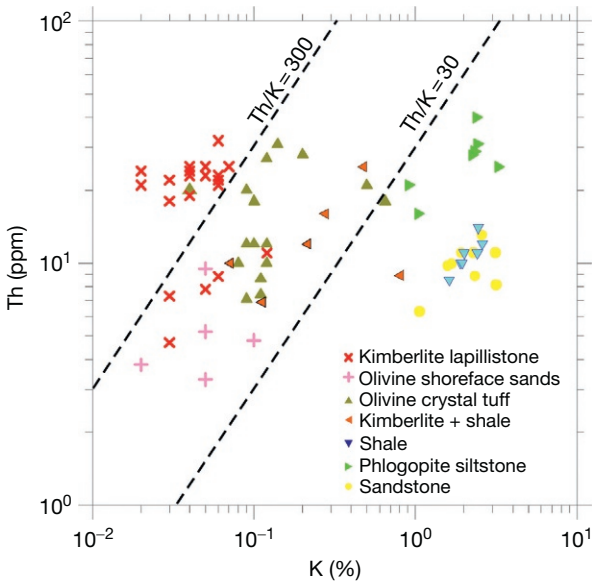


Figure 31 Relationship between logarithmic abundances of Th and K in Cretaceous sediments (shale, siltstone, and sandstone) and kimberlites. Dashed lines on the cross plot indicate constant Th/K ratios that separate three major groups (Mwenifumbo and Kjarsgaard, 2000). The cretaceous sediments (shale, sandstone, and phlogopite siltstone) cluster in the region of Th/K ratio <30 and the kimberlite cluster in the region of Th/K ratio >300. The Th/K ratio between 30 and 300 represents a mixture of sediments and kimberlitic material. Reproduced from Mwenifumbo CJ and Kjarsgaard BA (2000) Gamma-ray logging and radioelement distribution in the Fort a la Corne kimberlite pipe 169. Exploration and Mining Geology. *Journal of the Geological Society of the Canadian Institute of Mining* 8: 133–147.

individual radioactive elements – in particular, potassium, uranium, and thorium – in support of bedrock and surficial geologic mapping, exploration for a variety of mineral deposits, and environmental radiation monitoring. Directly or indirectly, measurements of variations in these three most common naturally occurring radioactive elements can provide information about the fundamental mineralogical and geochemical properties of bedrock and surficial deposits and can often indicate geologic or geochemical features not seen by other techniques.

Effective application of AGRS in geologic mapping and mineral exploration depends on two factors: (1) the extent to which radioelement distribution relates to bedrock differences, and the extent to which these are recognizably modified by mineralizing processes, and (2) the extent to which bedrock radioelement content is reflected in surficial materials that can be spatially related to bedrock sources.

Proper interpretation of AGRS data requires not only an understanding of the basic concepts of GRS as outlined in previous sections of this chapter but also a clear understanding of bedrock and surficial geology, petrology, and geochemistry of the radioactive elements.

The normal abundances and chemical properties of potassium, uranium, and thorium can provide insight into some of the fundamental mineralogical and geochemical properties of bedrock and surficial deposits and ore-related processes. Potassium, for example, is a major constituent of most rocks and occurs mainly in the mineral feldspar, which is a common and widespread mineral in the Earth's crust. It also occurs in other common rock-forming minerals, such as biotite and muscovite. Potassium concentrations in common rock-forming silicate minerals can range from 1% in hornblende, to 8% in biotite and muscovite, to 13% in orthoclase and microcline, and to 23% in nepheline (IAEA, 2003). The enrichment or depletion of potassium can also be indicative of alteration associated with a number of mineral deposit types including lode Au, VMS, alkaline and calc-alkaline porphyry Au–Cu (±Mo), and Olympic Dam-type Cu–Au–Fe oxides (Dickson and Scott, 1997; Ford et al., 2007; Shives et al., 1997).

Uranium and thorium, on the other hand, are generally present in low concentrations in most rock types (commonly measured in ppm) and occur in a wide range of minerals (IAEA, 2003). Uranium and thorium also represent relatively mobile and immobile elements, respectively, under most surface conditions, and are effective in the direct detection of some mineral deposits. Direct applications include the search for radioactive mineral deposits where uranium and thorium are a primary ore mineral, for example, unconformity-related U deposits, or where one or more of the radioactive elements are present as an associated trace element, for example, intrusion-hosted rare earth element (REE) deposits (Thomas et al., 2011). The range of radioactive element concentrations in most common rock types is shown in **Table 1** (Killeen, 1979).

Variations in radioelement ratio patterns may be particularly useful for mapping and mineral exploration applications as they can enhance subtle variations in elemental concentrations due to lithologic changes or alteration processes associated with mineralization. For example, potassium enrichment associated with certain types of mineral deposits may occur in host rocks with normally low to moderate levels of K, resulting in a high total K signature. This would be easily recognized if this alteration occurred in isolation. However, there may be normal, high K rock types in close proximity to the altered rocks. In this situation, the K associated with the alteration may not be distinguishable from other high K rock types. The ratio maps, in particular the eTh/K ratio, can be a sensitive indicator of K alteration associated with mineralization and can therefore be used as a direct indicator of mineralization. The addition of K associated with alteration will normally result in a low eTh/K ratio because thorium is relatively immobile and thus its concentration remains constant (Dickson and Scott, 1997; Ford et al., 2007; Portnov, 1987; Shives et al., 1997). Ford and O'Reilly (1985) have shown that on a regional scale, subtle variations in the measured variables eU and eTh are reinforced on the eU/eTh ratio map. These variations can be used to map cryptic (not visible) compositional variations in a suite of peraluminous granitic rocks where uranium concentrations generally increase and thorium concentrations decrease with increasing magmatic differentiation, resulting in abnormally high eU/eTh ratios associated with the most evolved parts of these granitic intrusions that have undergone late- and/or postmagmatic metasomatism and hydrothermal alteration. Identification of these more evolved and altered phases has exploration significance as granophile element mineralization is often associated with these phases (Ford and O'Reilly, 1985).

Given the various physical and chemical characteristics of the radioactive daughter products in the ^{238}U decay series, disequilibrium can be a source of error. Disequilibrium occurs

when one or more of the daughter products in the decay series are removed or added. In areas of extreme surficial weathering, disequilibrium is a significant concern (Dickson and Scott, 1997). In glaciated environments, disequilibrium due to weathering in bedrock is less of a concern but may be a concern if the effects of soil-forming processes are significant.

The effectiveness of GRS in geologic mapping and mineral exploration depends on several factors. As outlined earlier, these include the extent to which measurable differences in the radioactive element distributions relate to bedrock differences (normal lithologic signatures) and the extent to which these normal signatures are recognizably modified by mineralizing processes (K alteration). However, because AGRS is a surface technique only, the source of most gamma radiation is limited to the top 30–50 cm of the Earth's surface (Gregory and Horwood, 1961). This has profound implications for the interpretation of AGRS surveys. Whereas magnetic, electromagnetic, gravimetric, or seismic sensors may detect features to depths of tens or hundreds of meters, often buried far below the mappable near-surface geology, gamma ray data are interpretable in terms of the near-surface mineralogical and chemical composition of rocks and soils. This interpretation therefore requires an understanding of the nature of the surficial materials and their relationship to bedrock geology. For exploration applications, interpretation requires an understanding of the extent to which the ore or associated alteration signatures are incorporated into surficial materials, and whether these surficial materials can be spatially related to bedrock sources (Charbonneau et al., 1976).

Interpretation of AGRS survey data becomes complicated when bedrock-related signatures become obscured by overlying transported surficial material or by extreme weathering effects. In Canada, extensive glaciation has mechanically reworked and transported bedrock-derived material. However, the resulting overburden usually reflects local bedrock chemistry. Studies, for example, Shilts (1976), have shown that in glaciated terrains, the predominant component of glacial till is commonly locally derived, often within 1 or 2 km of its bedrock source. When the surficial material is composed predominantly of reworked material, such as glaciofluvial- or glaciolacustrine-type deposits, the resulting radioelement signature will not reflect the underlying bedrock. Therefore, glacial transportation of surficial material will often result in the radioelement signatures mapped at surface being somewhat displaced from their bedrock source and somewhat diluted in intensity. However, other studies, for example, Campbell et al. (2007) and Charbonneau et al. (1976), have shown situations where surficial material, with compositionally distinct radioelement signatures, has been transported considerable distances. Ford et al. (1988) and Thomas et al. (2011) documented two examples where glacial transportation of mineralized material, with compositionally distinct radioelement signatures at considerable distances from the source, significantly enlarged otherwise small exploration targets. Such examples clearly demonstrate the usefulness of AGRS surveys to support surficial mapping at regional and local scales. Therefore, effective interpretation requires some knowledge not only of the radioelement contents of rock types expected to be found in the survey area but also of the glacial and/or weathering effects in that area.

In Australia, Dickson and Scott (1997) documented the effects that weathering and regolith development have on normal, bedrock radioelement signatures. They reported that weathering of all rock types results in a loss of K and for felsic rock types in particular a reduction of U and Th as well. They reported typical losses of between 20% and 30% for K, U, and Th. For intermediate and basic rock types, they showed little change in radioelement concentrations during initial weathering, but with more advanced weathering and soil formation, U and Th concentrations may show a two- to threefold increase compared to unweathered equivalents. They emphasize that changes of these magnitudes can result in confusion when attempting to classify different rock types and their associated soils according to their typical radioelement signatures. They also reported that weathering and regolith development can significantly affect alteration signatures. Weathering of K alteration, particularly in the form of K-feldspar associated with some mineral deposits, will result in a loss of K, again making the identification of typical alteration signatures difficult. As in glaciated terrains where transported and reworked surficial material may mask underlying bedrock signatures, Dickson and Scott also noted that transported soils can mask, disguise, or change typical bedrock signatures often in unexpected ways.

This surface-only characteristic allows us to reliably relate the measured radioelement contrasts to mapped bedrock and surficial geology and alteration associated with certain mineral deposits. Unlike a camera lens, a gamma ray spectrometer does not have a 'fixed FOV' (see **Section 11.14.2.10.3**). The source of radiation may lie outside some nominal FOV. Therefore, a single measurement provides an estimate of the average surface concentration for an area of several thousand square meters. This single sample comprises variable proportions of exposed bedrock (fresh or weathered), overburden (glacial tills, glaciofluvial and glaciolacustrine, colluvium, alluvium, and loess), soils and regolith (clay, sand, and loam), soil moisture, standing water (lakes, rivers, swamps, and bogs), and vegetation. In almost all cases, the inhomogeneous nature of the surface (and therefore of each sample) introduces variability such that the absolute value of K, eU, or eTh may be less significant than the relative variations of the radioelement patterns.

The effects of variations in soil moisture, amount of bedrock exposure, or source geometry can be minimized through the use of radioelement ratios. In mineralized systems, all three radioelements may be enriched or depleted relative to unaltered host rocks. However, one or more of the radioelements may be preferentially affected for valid geochemical reasons. In these cases, the ratios offer very sensitive alteration vectors, which may lead directly or indirectly to the mineralization, even where the individual radioelement patterns are ambiguous.

GRS is particularly effective in granitic- and gneissic-dominated terrains where radioelement concentrations and contrasts are usually the highest (**Table 1**). However, depending on the flight line spacing and radioelement contrasts between different rock types, the technique is also effective in a variety of other terrains. When airborne surveys employ both a magnetometer and gamma ray spectrometer, the information obtained by the two techniques is often complementary, particularly in areas of complex geology.

11.14.8.3 Bedrock Geologic Mapping

In the 1970s, large systematic AGRS surveys were conducted in a number of countries (the United States, Canada, and

Sweden) to support national energy programs (Darnley and Ford, 1989). Initially, these large regional surveys were conducted with a coarse, reconnaissance scale line spacing. In Canada, these surveys were conducted under the Uranium Reconnaissance Program with a line spacing of 5000 m. In the United States, they were conducted under the National Uranium Resource Evaluation (NURE) program commonly with a line spacing of 8 miles. This coarse line spacing was chosen initially because very little information was known about the radioelement geochemical landscape for vast areas. Despite leading to the direct discovery of a number of mineral deposits (Shives et al., 1995), this coarse line spacing may actually have resulted in slower acceptance of AGRS for geologic mapping purposes. The coarseness of the reconnaissance surveys limits the detailed information often required to support modern geologic mapping activities. However, the point-to-point data collected along each flight line are detailed enough to support more detailed interpretation.

Ford et al. (2008a,b) presented an example from the southern Melville Peninsula, Quoich River area (**Figure 13**), Nunavut, where they demonstrated that despite a coarse line spacing of 5000 m, significant reconnaissance-level geologic information could be obtained from grids prepared using these data and that the detailed line data (**Figure 16(a)**) still supported local interpretation. The bedrock geology of the southern Melville Peninsula is dominated by a variety of older Archean gneisses and schists. The younger, Paleoproterozoic supracrustal Penrhyn Group metasediments consist of quartzites and marbles. Numerous aplitic to pegmatitic sills and plutons intrude the Penrhyn Group rocks. Many of these younger intrusions are anomalously radioactive and were exploration targets for intrusion-hosted uranium mineralization (Delpierre, 1982). Many of these younger intrusions and pegmatites can be identified by elevated eU and eU/eTh anomalies (**Figure 16(a)**) and by shades of blue on the ternary radioelement image (**Figure 13**) that represent a relative increase of eU with respect to K and eTh. The ternary radioelement image shows that subdivision of the various gneissic and supracrustal terrains and younger intrusive units is possible based on pronounced differences in their radioelement signatures.

Toward the southwest corner of the Quoich River area (**Figure 13**), the reconnaissance AGRS patterns show a prominent, generally northwest–southeast trending feature that truncates the otherwise dominant northeast–southwest trending regional structural and lithologic trends. This feature is characterized by generally lower radioelement signatures, reflected by the generally lower color intensity, that are particularly prominent in the eU and associated ratio maps (Ford et al., 2008a,b). On the ternary radioelement image (**Figure 13**), this feature is dominated by magenta colors that indicate a relative enrichment of K over eU and eTh. In the eastern part of this feature, near Wager Bay, Dredge et al. (2005) reported that the low AGRS signatures were coincident with an area dominated by meltwater erosional and depositional features. Interpretation of LANDSAT images elsewhere over this feature identified similar surficial characteristics (J. Harris, personal communication; Geological Survey of Canada, 2012). The scale of this unusual AGRS signature would suggest that these erosional and depositional features are much more extensive than originally documented. Identification and delineation of such surficial deposits by their distinct AGRS signatures would clearly play an important role for the interpretation of other regional surficial geochemical datasets including soil, till, and lake or stream sediment geochemistry.

AGRS data for the province of Nova Scotia (**Figure 14**) illustrate another example of the application of AGRS to mapping a mixed intrusive/metamorphic/sedimentary glaciated terrain. **Figure 14** shows the NADR, K, eU, eTh, eU/eTh, eTh/K, ternary radioelement image, and bedrock geology for the province of Nova Scotia (Keppie, 2000). The AGRS data were collected with a flight line spacing of 1000 m. The inset maps show AGRS data collected with a flight line spacing of 250 m for the area around the East Kemptville tin deposit in the southwestern part of the province. A variety of large- and small-scale features are evident to varying degrees in each map. For example, the K map distinguishes most of the granitic rocks and some internal compositional variations within these granitic rocks. In general, however, the restricted range of concentrations limits further internal subdivision in contrast to variations in eU and eTh concentrations. Subtle contrasting variations in K, eU, and eTh are highlighted on the eU/eTh and eTh/K ratio maps. Compared to the single variable and ratio maps, the ternary radioelement image appears to define many of the radioelement subdivisions. However, because the distribution of the K, eU, and eTh is histogram equalized to produce the ternary radioelement image, some information at the upper and lower limits of the data may be visually less apparent or lost altogether. Internal radioelement variations within the granite that hosts the East Kemptville tin deposit that are evident on the eU/eTh ratio map and point directly to the deposit area are not evident on the ternary radioelement image.

Follow-up studies (Ford, 1993; Ford and O'Reilly, 1985) in the East Kemptville area and elsewhere in Nova Scotia confirm a direct correlation between high AGRS eU/eTh ratio values and the most evolved phases of these composite granitic intrusions. These more evolved, high eU/eTh ratio phases exhibit other elemental variations that suggest varying degrees of late- and/or postmagmatic metasomatism and hydrothermal alteration. Granophile element mineralization, such as the East Kemptville tin deposit, is often associated with areas of alteration and anomalous radioelement signatures (Webster, 1984; Yeates et al., 1982). The common association of granophile element mineralization with areas of pervasive alteration and anomalous radioelement signatures has important applications in mineral exploration.

In addition to mapping and mineral exploration applications, the AGRS maps from around the East Kemptville area show an important surficial mapping application. Sharp increases in eU and eTh concentrations along the northwest contact of the host granite with the metasediments of the Meguma Group, southwest of the deposit, are coincident with an abrupt change between Meguma-derived tills to the northwest and granite-derived tills to the southeast. South of the deposit, along the southeast contact of the granite with the Meguma Group, radioelement patterns are more diffuse reflecting dispersion of granite-derived till southeastward over the Meguma Group bedrock for distances up to 5 km. In this example, the diminishing down-ice radioelement concentrations relate to diminishing down-ice concentrations of granite-derived clasts and matrix material.

Examples such as these tend to perpetuate the misconception that AGRS surveys are best suited for mapping exposed granite–gneissic-dominated terrains. Harris et al. (2008) have shown that AGRS can also provide important stratigraphic mapping assistance in mountainous, sedimentary-dominated terrains. **Figure 32** shows the geology, from Blusson (1971), K, eU, eTh, eU/eTh, and ternary radioelement images for the Sekwi Mountain area in the Mackenzie Mountains, along the border between the Northwest and Yukon Territories of Canada. The geology of the study area consists of a wide range of lithologic units including granites, shales, sandstones, and carbonates. Visual and computer-assisted classifications of the AGRS data indicated that many of these units could be successfully identified based on characteristic radioelement signatures. In a number of cases, distinct compositional subdivisions within some of the black shale and granitoid phases were also clearly mapped and confirmed by GGRS investigations (**Figure 33**).

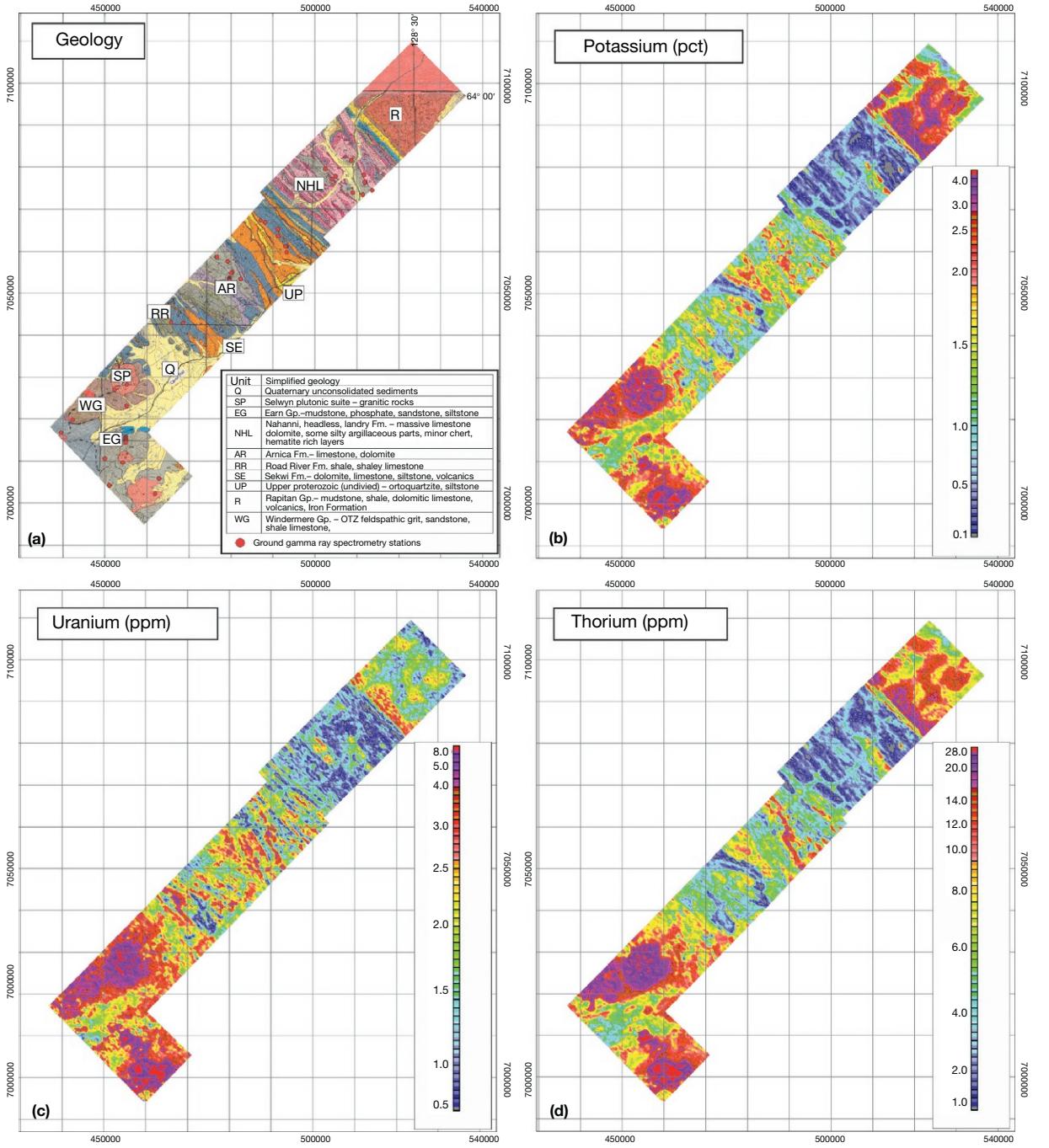


Figure 32 Images of AGRS survey data from the Sekwi Mountain area, Northwest Territories. (a) Geology of the Sekwi area (reproduced from Blusson SL (1971) *Sekwi Mountain Map Area, Yukon Territory and District of MacKenzie (NTS 105P)*, p. 17. Ottawa, ON: Geological Survey of Canada, Paper 71-71); (b) potassium (K, %); (c) equivalent uranium (eU, ppm); (d) equivalent thorium (eTh, ppm);

(Continued)

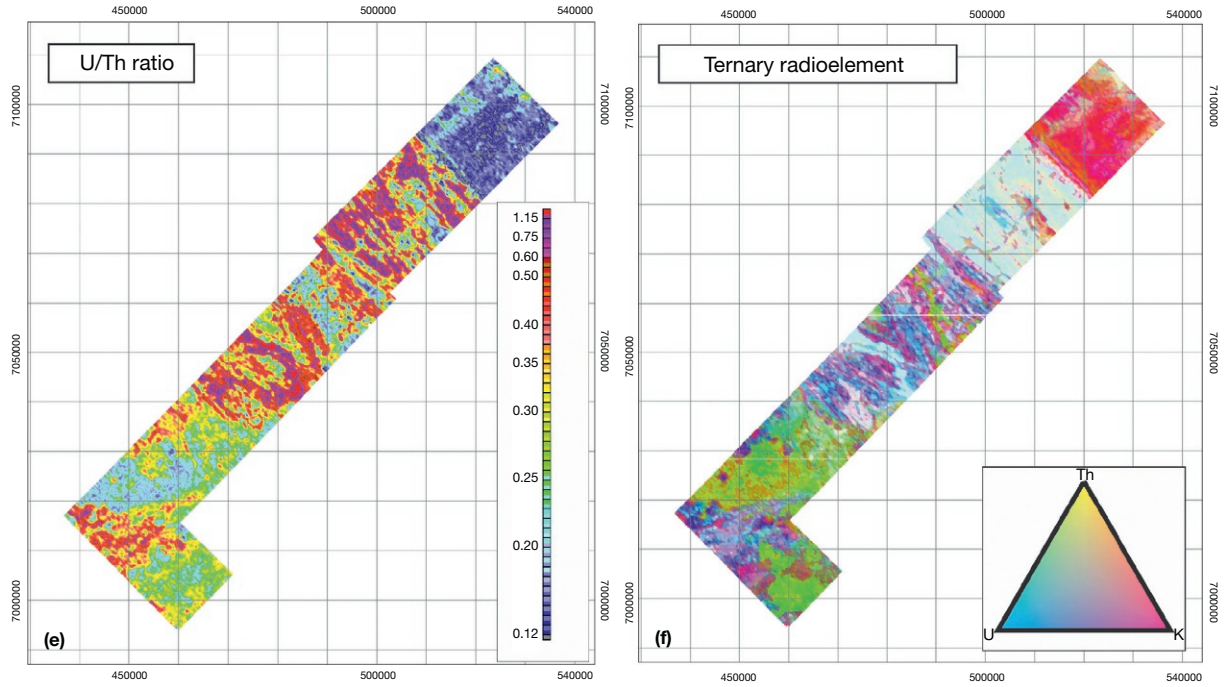


Figure 32 (Continued) (e) equivalent uranium/equivalent thorium ratio (eU/eTh); (f) ternary radioelement image (adapted from Harris JR, Ford K, Charbonneau B, and Buckle J (2008) Application of gamma ray spectrometer data for lithological mapping in a Cordilleran environment, Sekwi Region, Northwest Territories. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping, Case Study 5*, pp. 203–223. Ottawa, ON: Geological Survey of Canada, Open File 5643). Map grid is 10 km by 10 km.

11.14.8.4 Surficial Geologic Mapping

AGRS is a surface technique with the bulk of the measured gamma radiation originating from the top 50 cm of the surface. This, however, does not limit its usefulness as a bedrock mapping tool to areas of exposed bedrock or thin residual soils. AGRS measures variations in the bulk compositions of K, eU, and eTh of whatever material occurs at the surface. If this surface material is reflective of the underlying bedrock, then AGRS will be an effective bedrock mapping tool provided that sufficient radioelement contrasts exist in the different underlying bedrock units. In extreme situations where the surface material is composed almost exclusively of reworked or transported material, such as glaciofluvial, glaciolacustrine, alluvial, colluvial, or aeolian deposits, or in areas that have been subjected to varying degrees of in situ weathering, then its effectiveness as a bedrock mapping tool will be reduced or totally mitigated. However, if there is sufficient radioelement contrasts between the different surficial deposits, then AGRS will still be an effective surficial mapping tool.

Shilts (1976) noted that the main mass of most glacial tills is often locally derived. In these situations, the AGRS responses (similar perhaps to those in the East Kemptville area of southwestern Nova Scotia Figure 14) show patterns that are somewhat displaced and somewhat diluted in intensity compared to the bedrock source areas. However, if the surficial material is not reflective of the underlying bedrock, then AGRS responses similar perhaps to those shown for the southwestern part of the Quoich River area (Figure 13), where meltwater deposits mask and truncate distinctly different bedrock signatures, would be expected.

While radioelement domain mapping is often used to support bedrock geologic mapping, it can also be applied to support surficial geologic mapping. When an area is composed of two or more distinctly different, regional-scale, first-order radioelement signatures, the visual interpretation and classification of subtle radioelement variations within each domain can be improved by subsetting the dataset into separate, smaller datasets that better correspond to the distribution of each domain. Figure 34(a) shows the ternary radioelement map for a compilation of multiple AGRS surveys, flown with flight line spacings that range between 400 and 5000 m, over the Athabasca Basin and adjacent basement terrains of northern Saskatchewan. The various basement terrains clearly display more intense radioelement signatures compared to those over the sandstone-dominated rocks of the Athabasca Basin. Figure 34 (b) shows an enlargement of part of the same ternary image for the southeastern part of the Athabasca Basin and adjacent basement terrains. Figure 34(c) shows the same area; however, this image was generated using only AGRS data over the Geikie River map sheet. Figure 34(d) shows the same area; however, this image was generated using only AGRS data from the Athabasca Basin portion. By regenerating the ternary image using data from various subareas, the distribution of colors is better optimized to a smaller range of colors within that particular subarea. The visual contrasts were further

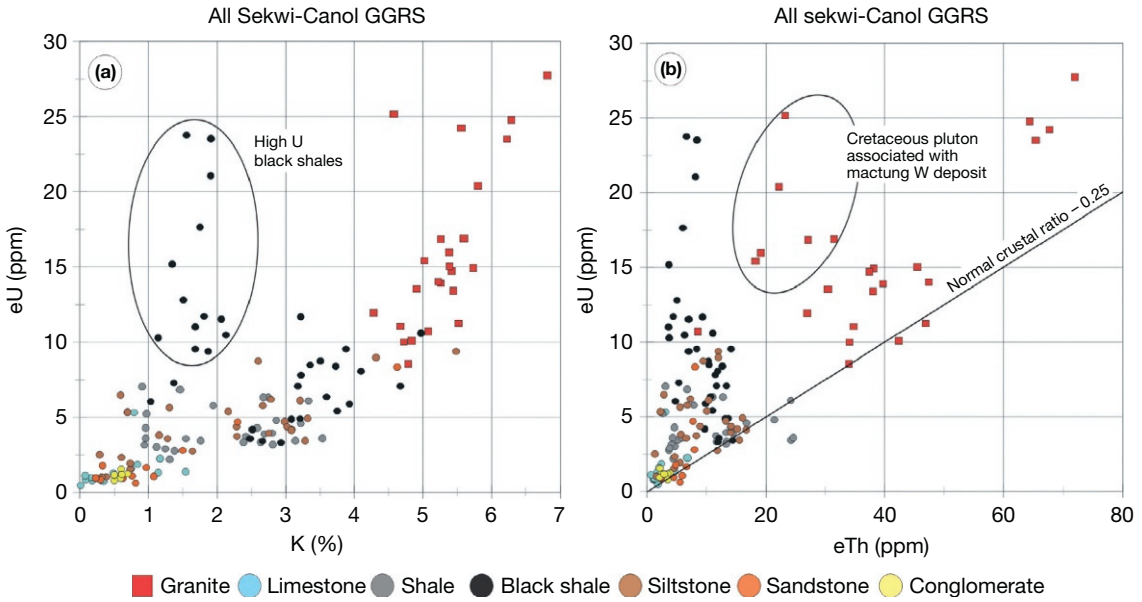


Figure 33 Cross plots showing the distribution of GGRS measurements color-coded by lithology. (a) eU (ppm) versus K (%). Note distinction of various sedimentary lithologies including subdivision of high eU black shale units. (b) eU (ppm) versus eTh (ppm). Note separation of high eU/eTh ratio values for pluton associated with MacTung deposit. Adapted from Harris JR, Ford K, Charbonneau B, and Buckle J (2008) Application of gamma ray spectrometer data for lithological mapping in a Cordilleran environment, Sekwi Region, Northwest Territories. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping, Case Study 5*, pp. 203–223. Ottawa, ON: Geological Survey of Canada, Open File 5643.

improved when data from over lakes and other wet or saturated ground were also removed (R. Fortin, personal communication; Geological Survey of Canada, 2013). Campbell et al. (2007) reported that defining radioelement domains from AGRS data over the Athabasca Basin provided a rapid means for mapping till compositions and for estimating regional variations in provenance. They also stated that separating areas of thin drift cover composed of local sandstone-derived tills from areas dominated by drift cover derived from basement sources would improve the effectiveness of other drift prospecting techniques such as till geochemistry.

Glacially derived material can reduce or enhance the exploration effectiveness of radioelement signatures measured by AGRS surveys. Ford et al. (1988) documented an example where glacial dispersion of material from a small, $<1\ km^2$, thorium-enriched but unexposed carbonatite resulted in a much larger target anomaly in excess of $10\ km^2$. Follow-up investigations that included GGRS and till geochemistry pinpointed the source carbonatite entirely masked by a small lake and documented a diminishing down-ice dispersion of eTh and REEs.

In areas where the effects of tropical weathering and landform development may significantly modify the original bedrock radioelement signatures, the AGRS patterns have been shown to provide important information for soil, regolith, and geomorphology studies for land use planning and mineral exploration (Wilford et al., 1997). Dickson and Scott (1997) reviewed results of studies conducted in Australia that showed how varying degrees of weathering modified the concentration and distribution of radioelements compared to those of unaltered bedrock. They noted that weathering led to a significant loss of K in all rock types and a loss of U and Th as well in felsic rocks, typically in the order of 20–30% for all three radioelements. In basic to intermediate rock types, extreme weathering and pedogenesis (soil-forming processes) can result in a two- to threefold increase in U and Th concentrations compared to unaltered bedrock. They reported that such modifications of radioelement signatures can lead to confusion in the classification of some AGRS signatures where soils derived from different rock types can have similar radioelement signatures.

11.14.8.5 Mineral Exploration

Variations in the abundances and chemical properties of K, U, and Th can provide insight into ore-related processes and may be a direct or indirect aid to mineral exploration for a variety of commodities. For example, K is a major constituent of most rocks and is also diagnostic of alteration associated with many mineral deposits. Uranium and thorium on the other hand are usually present in trace amounts, measured in ppm, and are relatively mobile and immobile elements, respectively, under most surface conditions. They are effective in the direct detection of mineralization if they are the primary target, for example, uranium deposits, or if they are an associated trace element, for example, intrusion-hosted REE deposits.

Dickson and Scott (1997) provided a review of the application of AGRS to exploration for a variety of mineral deposit types in Australia, including granophile Sn, W, and Mo; porphyry Cu–Au; Au; and stratabound polymetallic deposits. They noted that radioelement distributions associated with each type of deposit are varied and complex. An understanding of the effects of K alteration, varying degrees of in situ weathering,

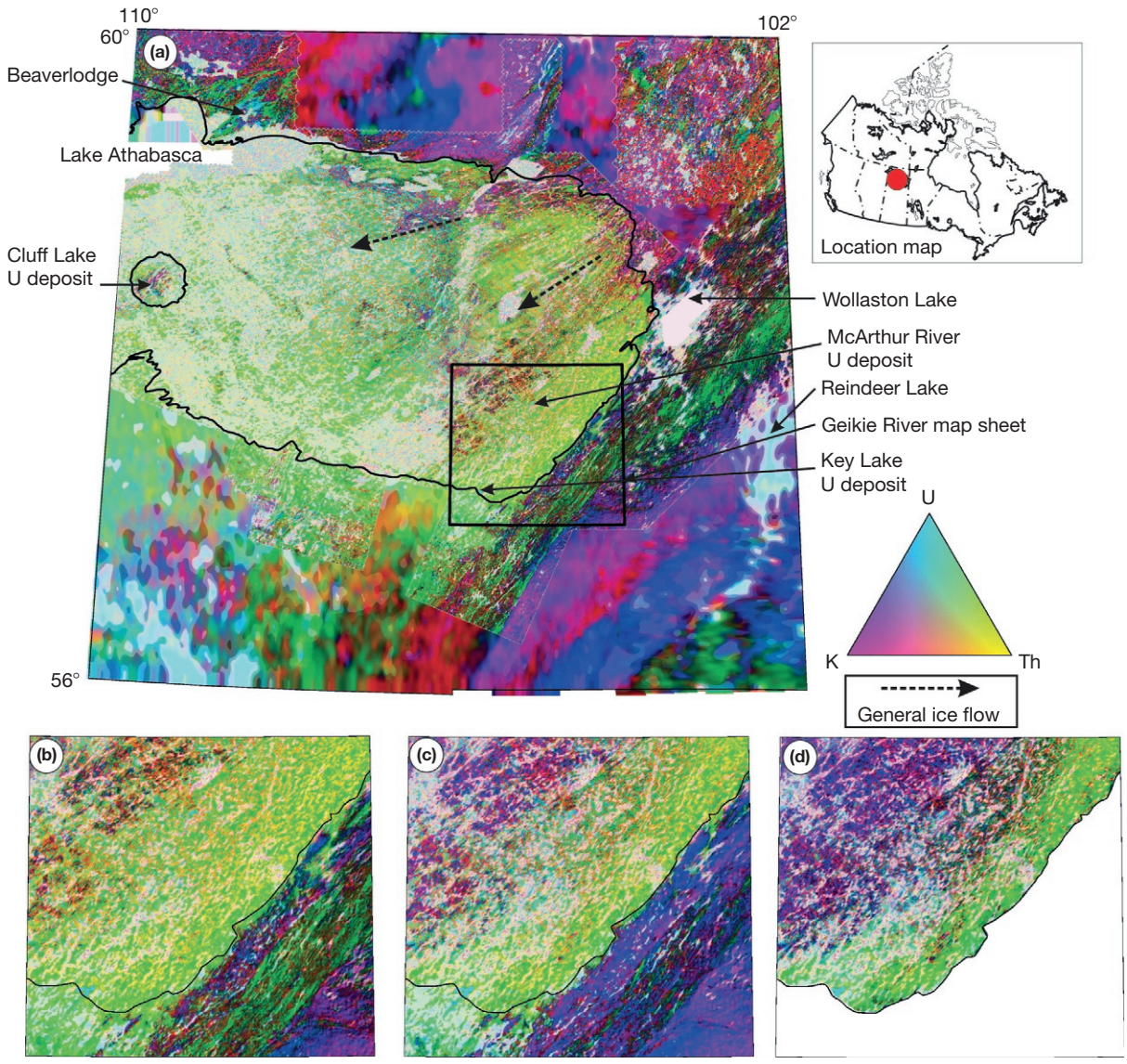


Figure 34 Series of ternary radioelement images compiled from multiple AGRS surveys, flown with flight line spacings that range between 400 and 5000 m over the Athabasca Basin and adjacent basement terrains, northern Saskatchewan (Buckle et al., 2011). Contact between Athabasca Basin rocks and basement rocks shown by irregular black line. (a) Ternary radioelement image, northern Saskatchewan. (b) Same ternary radioelement image for Geikie River area as shown in (a). (c) Ternary radioelement image generated using only AGRS data for Geikie River area. (d) Ternary radioelement image generated using only AGRS data for Athabasca Basin portion of Geikie River area. Note change in radioelement domains for different subsets of AGRS data used (images provided by R. Fortin, Geological Survey of Canada). Reproduced from Buckle JL, Carson JM, Miles WF, Ford KL, Fortin R, and Delaney G (2011) Geophysical Series, Geophysical Compilation Northern Saskatchewan. Geological Survey of Canada, Open File 7039; Saskatchewan Ministry of Energy and Resources, Open File 2011-58; scale 1:750 000.

and local lithologic variations on these radioelement distributions is required to properly evaluate the effectiveness of AGRS as an exploration technique.

11.14.8.5.1 Direct detection: uranium

Probably, the most direct application of AGRS in mineral exploration is for uranium. This direct application has been described in many publications, for example, IAEA (1976, 1979), Killeen (1979), and Darnley et al. (1977). Potentially, economic uranium mineralization is often characterized by abnormally high eU/eTh ratio values that can be easily identified in stacked profiles (**Figure 16**) and gridded presentations. Proper ratioing of eU with eTh and/or K data can often enhance subtle variations in uranium concentrations. However, caution should be exercised when interpreting ratio patterns as high ratio anomalies, with no exploration significance, can be generated in low count rate areas (Shives et al., 1995). **Figure 35** shows an example of AGRS data for eU (ppm), eTh (ppm), and eU/eTh over pegmatite-style uranium mineralization from the Bancroft area, Ontario. This area is well known for abundant U- and Th-bearing pegmatites, veins, and skarn-type deposits shown by the black dots in **Figure 35(a)** and **35(b)**, several of which have been mined in the past. Clearly, the majority of mineralized occurrences are directly associated with or occur in close proximity to anomalous AGRS signatures. The eU/eTh map shows the main zone

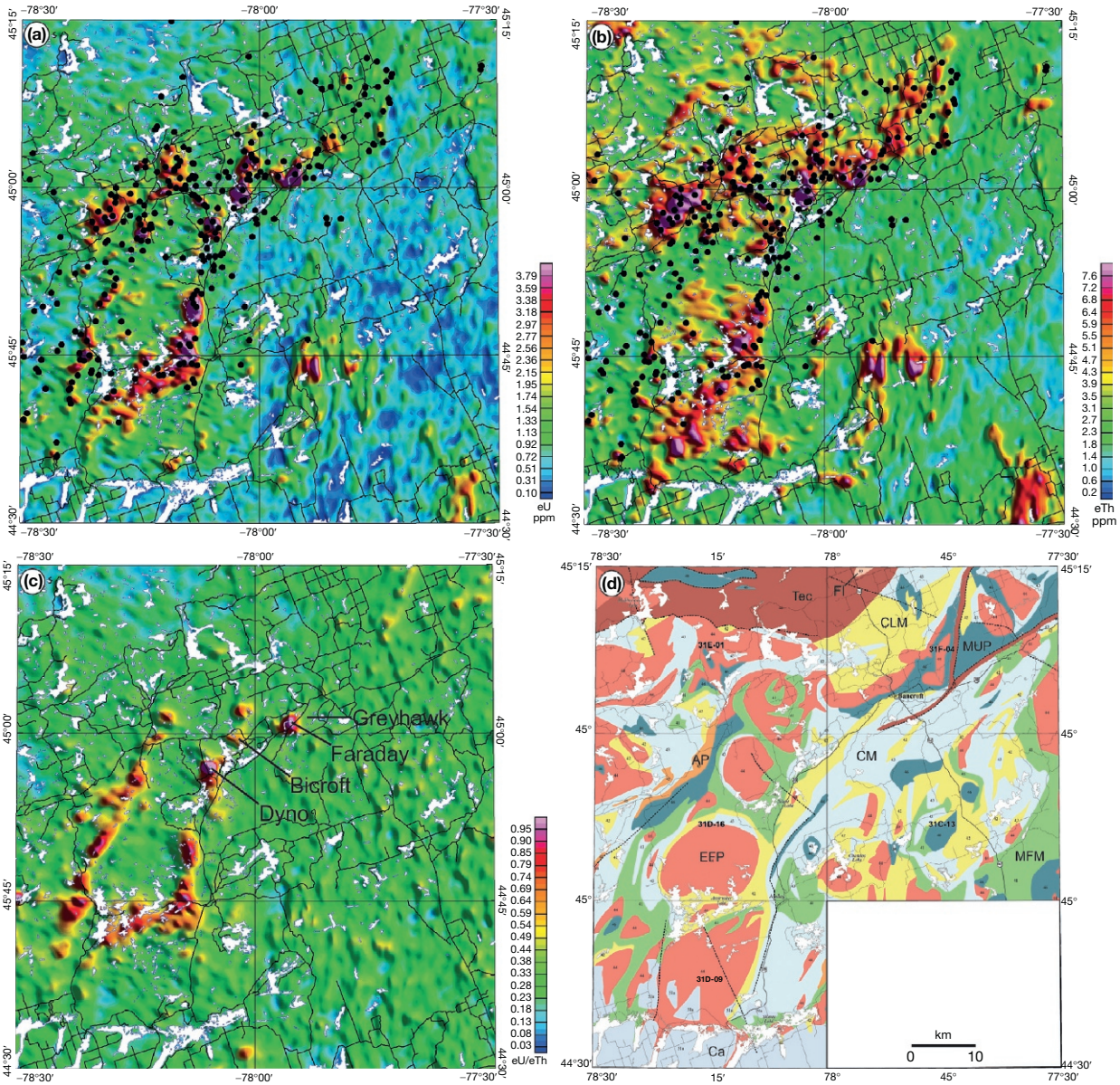


Figure 35 Airborne gamma ray spectrometric maps for (a) eU (ppm), (b) eTh (ppm), and (c) eU/eTh compiled from 1000 m line spaced data, and (d) bedrock geology (Ontario Geological Survey, 1991) for the Bancroft area, eastern Ontario. Black dots show locations of all known radioactive mineral occurrences including past producing mines at Faraday, Bicroft, Dyno, and Greyhawk. Geology panel provided by R. McNeil, Geological Survey of Canada. Ca, paleozoic carbonate; Tec, tectonite; MUP, mafic to ultramafic plutonic; AP, alkalic plutonic; EFP, early felsic plutonic; CM, carbonate metasedimentary; CLM, clastic metasedimentary; MFM, mafic to felsic metasedimentary; FI, felsic igneous. Adapted from Ontario Geological Survey (1991) Bedrock geology of Ontario, southern sheet. Ontario Geological Survey, Map 2544, scale 1:1 000 000.

of pegmatite mineralization. Those that show an enrichment of eU over eTh, above the normal U/Th crustal ratio of 0.25 (Clark et al., 1966), are structurally distributed around the margins of a series of older gneissic units.

11.14.8.5.2 Direct detection: U and Th as trace elements

U and Th often serve as pathfinder or indicator elements for a number of other commodities, including REEs, Li, Cs, Be, Nb, Ta, and Zr, which are concentrated in some types of pegmatites, granitic intrusions, carbonatites, and alkalic complexes. In these situations, the U and Th concentrations may not be sufficiently high to warrant extraction but occur as associated trace elements at sufficient concentrations to produce a distinctive radioelement signature. Thomas et al. (2011) reviewed geophysical methods, including AGRS that have proven effective in the search for intrusion-hosted REE deposits. AGRS examples include complex pegmatites (Bernic Lake, Manitoba), alkaline–peralkaline intrusions (Strange Lake, Quebec/Labrador, and Thor Lake, Northwest Territories), and carbonatites (Allan Lake, Ontario).

Figure 36 illustrates a classic example of the effectiveness of AGRS for targeting a REE-enriched carbonatite. Although this carbonatite represents a small, unexposed (hidden beneath a lake and glacial overburden) target, glacial dispersion of

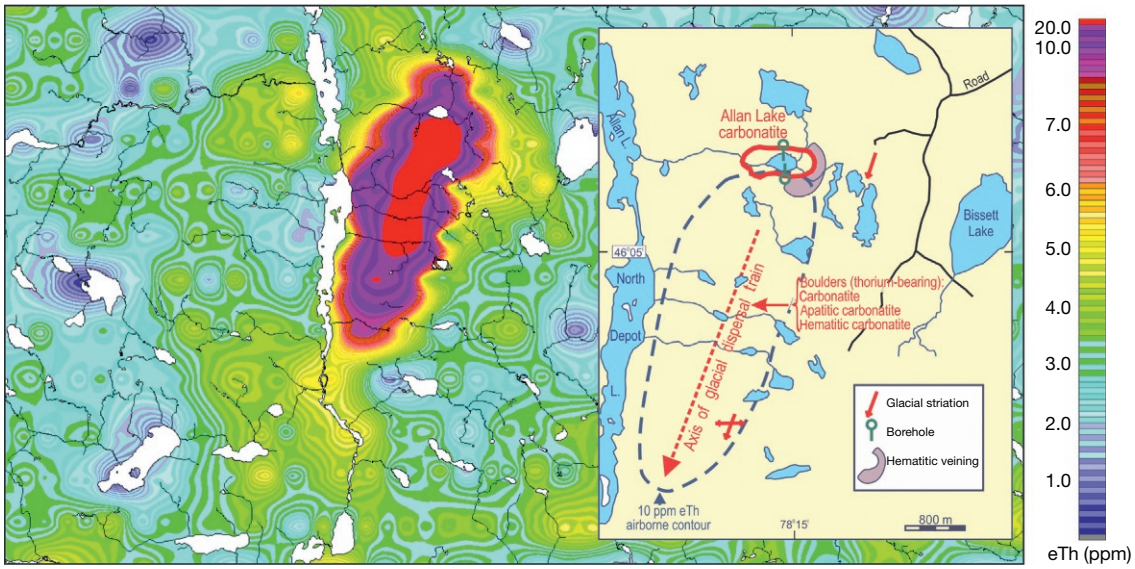


Figure 36 Airborne gamma ray spectrometric map for eTh (ppm) compiled from 500 m line spaced data over a small, unexposed rare earth-enriched carbonatite (see inset map). Note significant enlargement of anomaly due to glacial dispersion.

Th-enriched carbonatite material produced a large dispersal train characterized by a distinctive eTh AGRS anomaly. The initial discovery of this dispersal train by a reconnaissance, 5000 m line spacing AGRS survey illustrates the effectiveness of AGRS as a regional exploration tool for REEs through the use of radioelements as pathfinder or indicator elements. A subsequent, detailed, 500 m line spacing AGRS survey clearly outlined the extent of glacial dispersion and pinpointed the source of the anomalous material. Follow-up investigations (Ford et al., 1988) that included GGRS, till geochemistry, and boulder mapping further defined the dispersal train and indicated the anomalous, multielement (Ba, Nb, Ce, La, Zn, Mn, and Fe) nature of the till derived from the carbonatite. This example demonstrates the utility of AGRS in a glaciated environment, where glaciation has produced a larger anomaly than would normally be observed in a nonglaciated environment.

The Strange Lake alkalic intrusion on the Quebec/Labrador border represents another example of the potential of the AGRS method for targeting intrusion-hosted REE mineralization. Postdiscovery AGRS surveys flown by the Geological Survey of Canada (1980, 1985) outlined anomalous eU and eTh (**Figure 37**) signatures that extended for approximately 50 km down-ice from the intrusion. The Allan Lake and Strange Lake examples represent the targeting of glacially transported anomalies that can be traced back to mineralization. The discovery of the Thor Lake REE deposit, hosted by the Blatchford Lake alkaline–peralkaline intrusive complex in the Northwest Territories, illustrates the direct detection of REE mineralization. In this case follow-up, investigations of prominent eU and eTh radioelement anomalies defined by a reconnaissance AGRS survey led to the discovery of multiple zones of mineralization (Charbonneau and Legault, 1994).

The effectiveness of AGRS to map granophile element mineralization within granitic rocks has been described by a number of authors in Australia and Canada (Ford, 1993; Ford and O'Reilly, 1985; Webster, 1984; Yeates et al., 1982). **Figure 14** shows the radioelement distribution patterns for 1000 and 250 m line spacing AGRS data over an area around the East Kemptville Sn deposit in southwestern Nova Scotia. The ternary radioelement image subdivides various phases of the Davis Lake pluton that hosts the deposit. However, it is the eU/eTh ratio map that provides a more direct vector to the mineralization.

11.14.8.5.3 Indirect detection: alteration mapping

Potassium alteration, at either a district or deposit scale, commonly in the form of sericite or potassium feldspar, is associated with many types of mineral deposits including VMS and Au deposits (Dubé et al., 2007; Galley et al., 2007) and shear-hosted Au deposits (Dubé and Gosselin, 2007). Many alkaline and calc-alkaline porphyry Cu–Au deposits have extensive potassic alteration halos (Davis and Guilbert, 1973; Sinclair, 2007). Potassium alteration in the form of biotite, K-feldspar, and sericite is also diagnostic of polymetallic, iron oxide Cu–Au deposits (Corriveau, 2007). Because Th does not often accompany K during hydrothermal alteration, the enrichment of K can often be distinguished from high K associated with normal lithologic signatures by virtue of a characteristic low eTh/K ratio (Shives et al., 1997).

11.14.8.5.3.1 Gold deposits

In the Red Lake area, northern Ontario, K alteration associated with Au mineralization is known to occur along the Madsen–Starrett-Olsen shear zone. Durocher (1983) noted that this alteration can result in a two- to fourfold increase in K concentrations in the volcanic rocks that host the Au mineralization relative to unaltered equivalent lithologies. **Figure 38** shows a weak but distinctive AGRS K anomaly associated with the mineralized shear zone. This is considered unusual given that the mafic volcanics, which host the mineralization, would be expected to have a very low K signature. Prominent K anomalies occur elsewhere in the survey area associated with granitic and felsic volcanic lithologies. The K enrichment associated with the mineralized zones can be distinguished from these other high

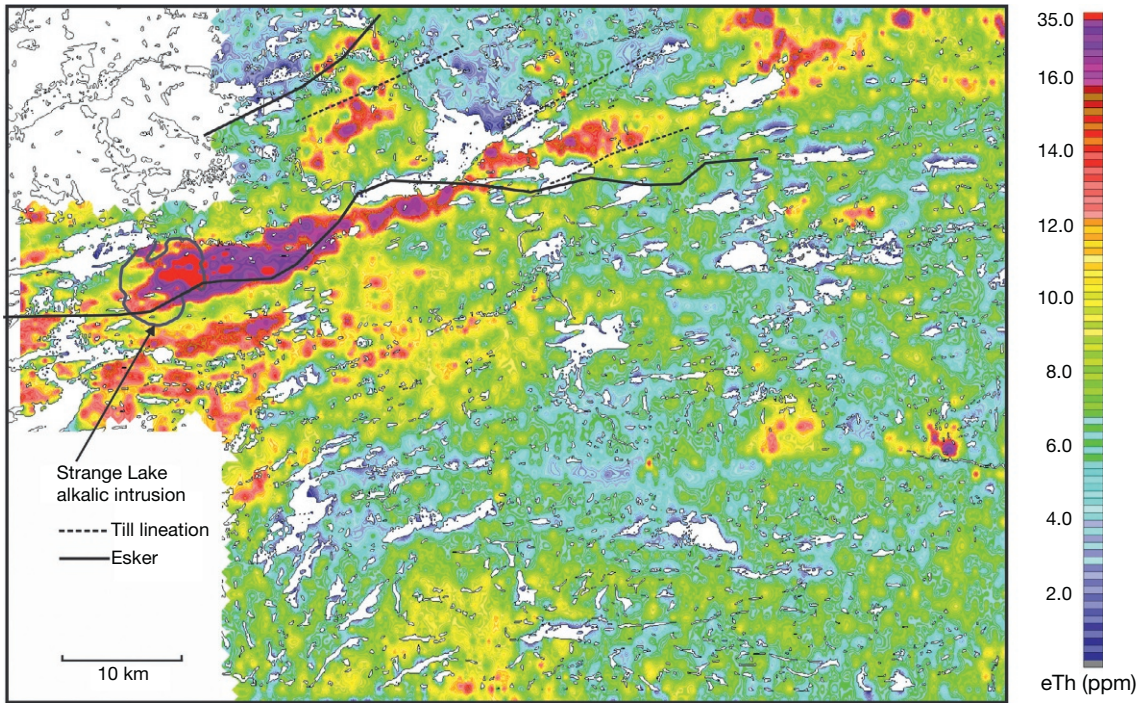


Figure 37 Airborne gamma ray spectrometric map for eTh (ppm) compiled from 1000 m line spaced data over the Strange Lake alkalic intrusion. Note delineation of extensive eTh-enriched till down-ice from the mineralized intrusion.

K anomalies by a coincident low eTh/K ratio anomaly (**Figure 38**) resulting from the fact that the mafic volcanics that host the mineralization have low Th concentrations. Ground follow-up studies using GGRS measured K concentrations consistent with those reported by Durocher (1983) for altered and unaltered lithologies (Shives et al., 1995).

Similar high K and low eTh/K ratio AGRS signatures were reported by Shives et al. (1995) for the Hemlo Au deposit in northwestern Ontario. Despite significant surface disturbance due to mining operations, measurements made by GGRS confirmed well-defined high K and low eTh/K ratio anomalies. Potassium concentrations increase from generally $<2\%$ K in unaltered rocks to concentrations in excess of 8% K in equivalent mineralized rocks.

11.14.8.5.3.2 Polymetallic IOCG-type deposits

Several authors have reviewed the geophysical signatures associated with polymetallic, Olympic Dam-type ($Fe \pm Cu \pm U \pm Au \pm REE$) deposits (Ford et al., 2007; Smith, 2002). In addition to regional and localized aeromagnetic anomalies, potassic alteration can be a diagnostic alteration signature. These alteration signatures may be large and regionally extensive, providing a significant exploration target for combined AGRS and aeromagnetic surveys. Because U mineralization is commonly associated with these deposits, high eU and eU/eTh AGRS signatures would also be expected.

The southern part of the Great Bear Magmatic Zone (GBmz) in the Northwest Territories hosts several IOCG-type deposits (Corriveau, 2007). The most prominent of these are the Sue-Dianne and NICO deposits. Follow-up investigations of strong eU and eU/eTh anomalies from AGRS reconnaissance surveys with 5000 m line spacing conducted by the GSC in the 1970s led to the discovery of the Sue-Dianne Cu–U–Au deposit (Charbonneau, 1988). In 1993, the GSC conducted a combined AGRS and magnetic total field survey over part of the southern GBmz to support ongoing geologic mapping and metallogenic studies (Hetu et al., 1994). **Figure 39** shows some of the AGRS and total field magnetic data for a portion of this detailed airborne survey over the NICO deposit. These images show a large K anomaly coincident with a positive magnetic total field anomaly. The prominent, low eTh/K ratio anomaly distinguishes the high K anomalies related to the extensive hydrothermal alteration system from other high K anomalies associated with normal lithologic variations. These combined AGRS and aeromagnetic survey results place several previously known, but noncorrelated mineral occurrences in the area into the context of a large, K- and Fe-enriched, polymetallic hydrothermal alteration system (Shives et al., 1997).

11.14.8.5.3.3 VMS deposits

A review of geophysical signatures associated with Canadian ore deposits was presented by Ford et al. (2007). In it, they qualitatively rate the AGRS method as moderately effective for direct targeting of VMS deposits. **Figure 17** modified from Shives et al. (1997) shows AGRS K data and GGRS follow-up results from the Pilley's Island VMS deposit, Newfoundland. The geology of the Pilley's Island area is dominated by tholeiitic and calc-alkaline mafic volcanics with localized felsic volcanic centers. AGRS K anomalies are commonly associated with these felsic volcanic centers. **Figure 17(a)** shows two of the strongest K anomalies associated with hydrothermally altered felsic pyroclastics and massive dacitic flows. **Figure 17(b)** and **17(c)** shows a wide range of K concentrations measured by GGRS that reflect varying degrees of potassic alteration. Locally,

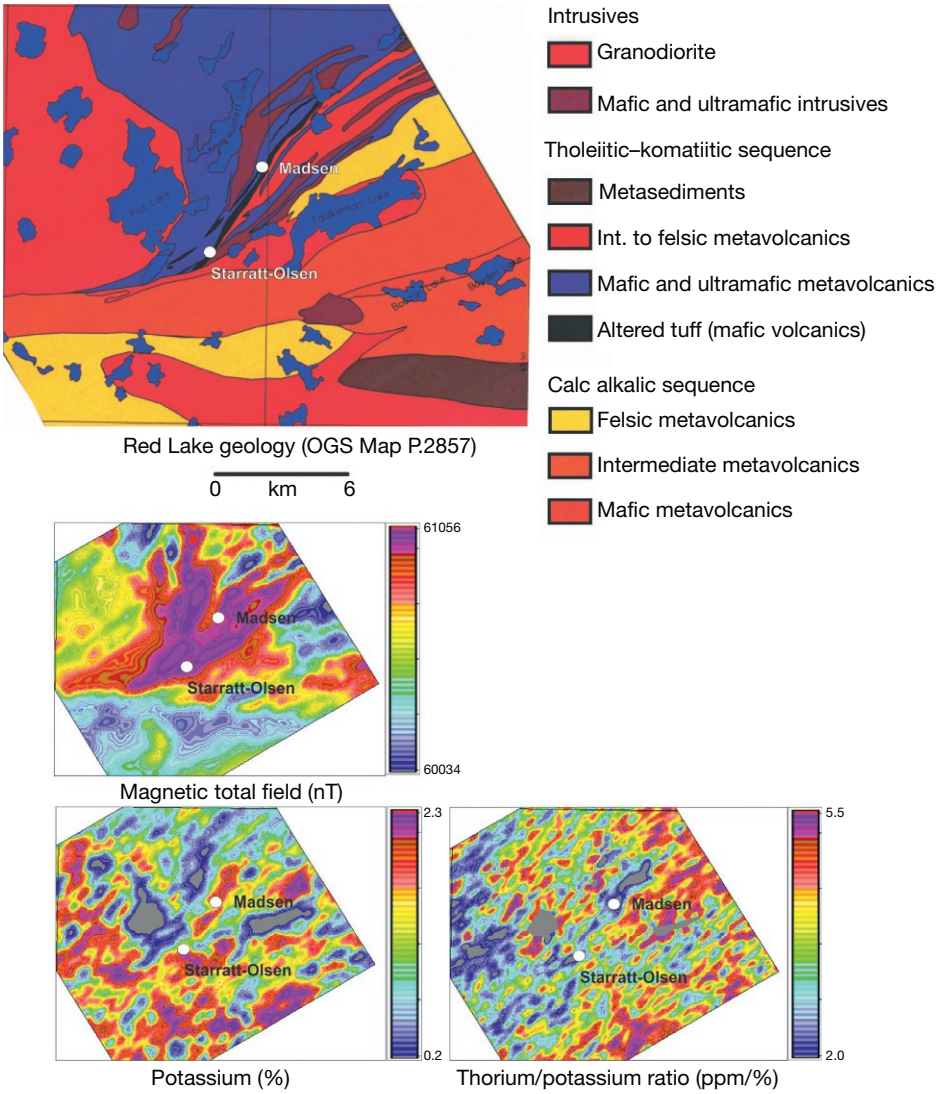


Figure 38 Airborne gamma ray spectrometric and magnetic signatures over gold mineralization in the Madsen–Starrett-Olsen area, Red Lake, Ontario. In this example, structural information is provided by the aeromagnetic data, lithologic information is provided by both the aeromagnetic and AGRS data, and K enrichment due to alteration associated with mineralization is differentiated from high K lithologies by corresponding low eTh/K ratios. White dots represent known gold mineralization. Adapted from Ford KL, Keating P, and Thomas MD (2007) Overview of geophysical signatures associated with Canadian ore deposits. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 939–970. St. John's, NL: Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

maximum K concentrations represent a two- to threefold increase over average K concentrations for unaltered to weakly altered felsic volcanic rocks of the Robert's Arm Group and a fourfold increase in K concentrations for altered pillow basalts at Bumble Bee Bight compared to unaltered mafic volcanics of the Robert's Arm Group. The ability to map and quantify this potassic alteration from the air and on the ground, and separate it from other normal, high K lithologic signatures, has important implications for VMS exploration in this and similar environments.

11.14.8.5.3.4 Porphyry deposits

Potassic alteration, in the form of a proximal biotite and/or K-feldspar alteration followed by a more distal, but often overlapping phyllic alteration that includes sericite, is often associated with porphyry Cu and Cu–Au deposits (Sinclair, 2007). The AGRS signatures of several porphyry deposits in western Canada have been presented by Shives et al. (1997). The potassium and total field aeromagnetic signature, compiled from a 500 m line spacing geophysical survey over the Mount Milligan porphyry deposit, is shown in **Figure 40**. A large, regionally extensive total field magnetic anomaly is associated with exposed and buried parts of the Mount Milligan intrusive complex. High K values are associated with bedrock exposures of the intrusive complex at Mount Milligan. To the south, discrete K anomalies associated with the Mount Milligan deposits provide smaller, more focused targets compared to the regional magnetic anomalies. Elsewhere in the

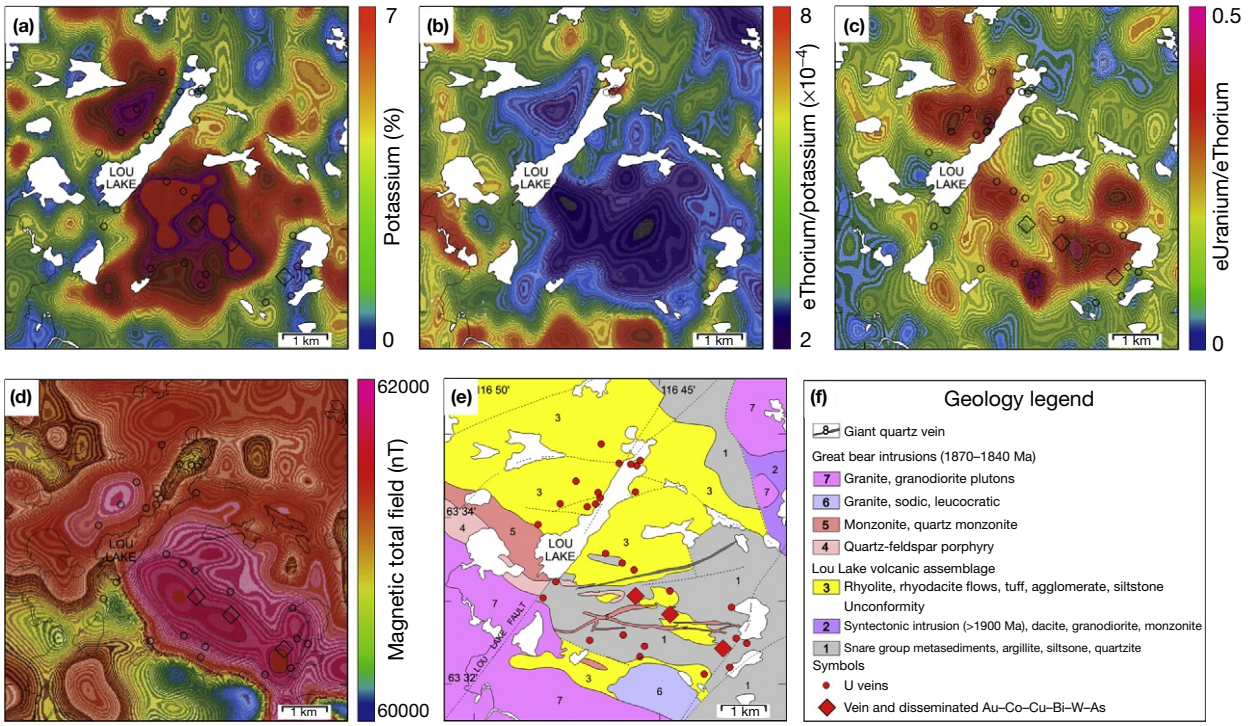


Figure 39 Airborne gamma ray spectrometric and total field magnetic maps over polymetallic IOCG-type mineralization near Mazenod Lake, Northwest Territories (Shives, Charbonneau and Ford, 1997). (a) Potassium (%), (b) eTh/K ratio, (c) eU/eTh ratio, (d) total magnetic field, (e) bedrock geology map, and (f) geology legend. Reproduced from Shives RBK, Charbonneau BW, and Ford KL (1997) The detection of potassic alteration by gamma ray spectrometry – Recognition of alteration related to mineralization. In Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial Conference on Mineral Exploration*, pp 741–752. Canada: GEO F/X.

survey area, despite few bedrock exposures and extensive, often thick till, glaciofluvial, and colluvial deposits, other discrete AGRS K anomalies (K5 and K6, **Figure 40**) are associated, respectively, with K-altered and mineralized andesitic volcanics (K5) and K-altered and mineralized intrusive boulders (K6) derived from an unexposed mineralized intrusive unit (Ford et al., 2007; Shives et al., 1997).

AGRS signatures over the Casino Au–Cu–Mo porphyry deposit in west-central Yukon Territory (**Figure 41**) show a high K anomaly associated with the Casino deposit. A unique, low eTh/K ratio anomaly separates high K values associated with altered and mineralized bedrock at the Casino deposit, from similarly high K signatures associated with felsic volcanic rocks in the area. The Casino deposit is unique in Canada because it occurs in an unglaciated, deeply weathered area. Weathering effects extend to 300 m below surface (Shives et al., 1997). Despite this deep weathering, the high K, low eTh/K ratio signature is still well defined by both AGRS and GGRS.

In addition to the several examples presented here that show the effectiveness of AGRS and GGRS for the direct and indirect detection of mineralization and associated alteration, Shives et al. (1995) had also shown the AGRS method to be effective for exploration for a number of other mineral deposit types in Canada. These include volcanic-hosted, epithermal base, and precious metals at Jacquet River area, New Brunswick; metasediment-hosted skarn mineralization (Au–U–W–Mo–Cu) at Sandybeach Lake, Nunavut Territory (Charbonneau and Swettenham, 1986); dolomite-hosted epigenetic U–Cu at South March, Ontario; limestone-hosted U at South Maitland, Nova Scotia (Charbonneau and Ford, 1979); and granite-hosted Au mineralization associated with illite alteration at Cape Spencer, New Brunswick.

11.14.8.6 Interpretation and Applications of BGRS

11.14.8.6.1 Geologic mapping: definition of geologic contacts

As mentioned earlier, natural gamma radiation from potassium, uranium, and thorium can be used to characterize rock types since the distribution and concentration of these naturally occurring radioelements vary between different rock types. Natural gamma ray logging can provide an important tool for mapping lithology and for stratigraphic correlation. The main minerals contributing to an increase in K are K-feldspars, micas, and clay minerals, such as illite. Uranium and thorium generally occur in accessory minerals such as allanite, apatite, monazite, sphene, xenotime, and zircon (Adams and Weaver, 1958; Serra et al., 1980). Refer also to **Section 11.14.8.2**.

The literature has numerous examples of applications of gamma ray logs in sedimentary geologic environments (primarily in the coal, oil, and gas industry). They include lithologic mapping, stratigraphic correlation, and the determination of amount and type of clay minerals present in clastic sediments. Information on the depositional environments, diagenesis, and provenance of the sediments was often extracted from the

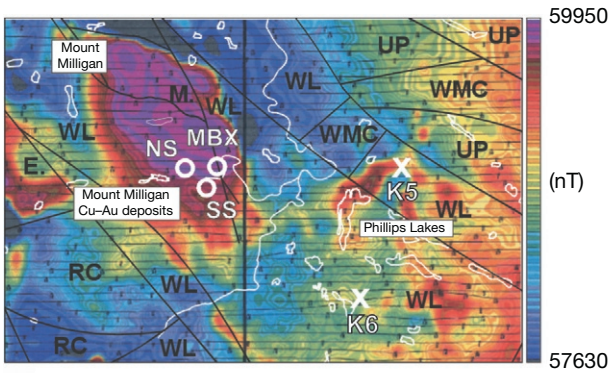


Figure 40 Magnetic total field (nT) and AGRS K (%) signatures in the area of the Mount Milligan porphyry copper deposits (MBX, NS, and SS), British Columbia. E, volcanic wacke; M, Mount Milligan intrusive complex, monzonite suite; RC, rainbow creek formation; UP, mesozoic sedimentary rocks; WL, Witch Lake formation (Takla Group); WMC, Wolverine metamorphic complex. Reproduced from Shives RBK, Charbonneau BW, and Ford KL (1997) The detection of potassic alteration by gamma ray spectrometry – Recognition of alteration related to mineralization. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial Conference on Mineral Exploration*, pp. 741–752. Canada: GEO F/X.

distribution of the radioelements (Adams and Weaver, 1958; Fertl, 1979; Serra et al., 1980).

The use of gamma ray logs for mapping lithology in hard rock mining environments is not as well established as in the sedimentary rock environments in the petroleum and coal industry. However, the visual identification of different metamorphic and igneous rocks can be greatly improved with the aid of gamma ray measurements. Accurate identification of different ultramafic and mafic volcanic units can improve the understanding of the geology and structure in mineral exploration camps. Mwenifumbo and Killeen (1986) used TC gamma ray logs to characterize different volcanic units at a mining camp in Newfoundland, Canada. Figure 42 shows how the TC gamma ray log was useful in mapping a sequence of volcanic rocks at the Buchans Mine in Newfoundland. The rhyolite, which contains the most potassic feldspars, shows higher gamma ray activity than the diabase, which contains the lowest feldspar content, and the andesite, which has intermediate content. Lead, zinc, and copper sulfide mineralizations are closely associated with rhyolitic volcanics.

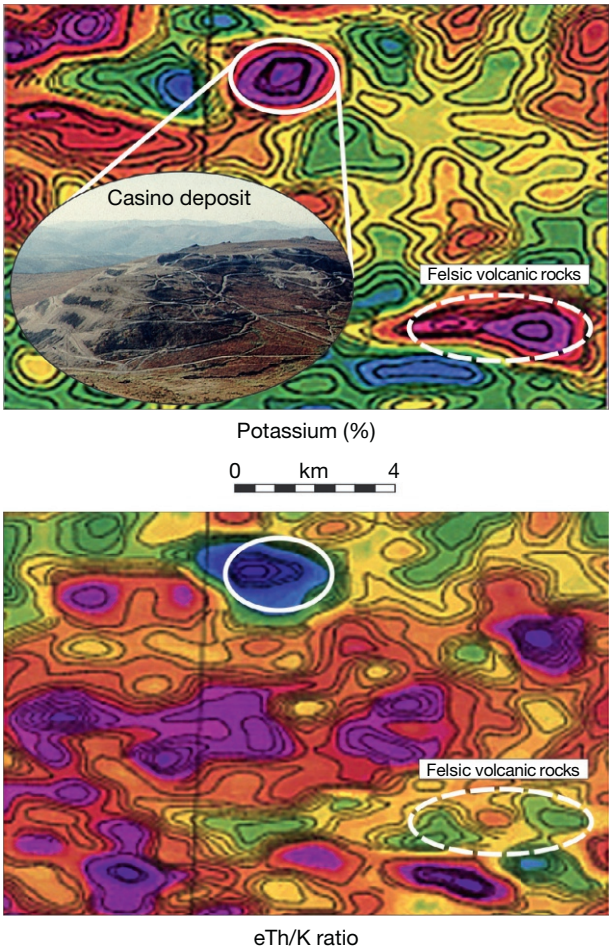


Figure 41 Airborne gamma ray spectrometric maps for K (%) and eTh/K ratio compiled from 500 m line spaced data over the Casino deposit area, and line data displayed as a stacked profile showing K enrichment and low eTh/K ratio values characteristic of K alteration associated with porphyry mineralization. Reproduced from Shives RBK, Charbonneau BW, and Ford KL (1997) The detection of potassic alteration by gamma ray spectrometry – Recognition of alteration related to mineralization. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial Conference on Mineral Exploration*, pp. 741–752. Canada: GEO F/X.

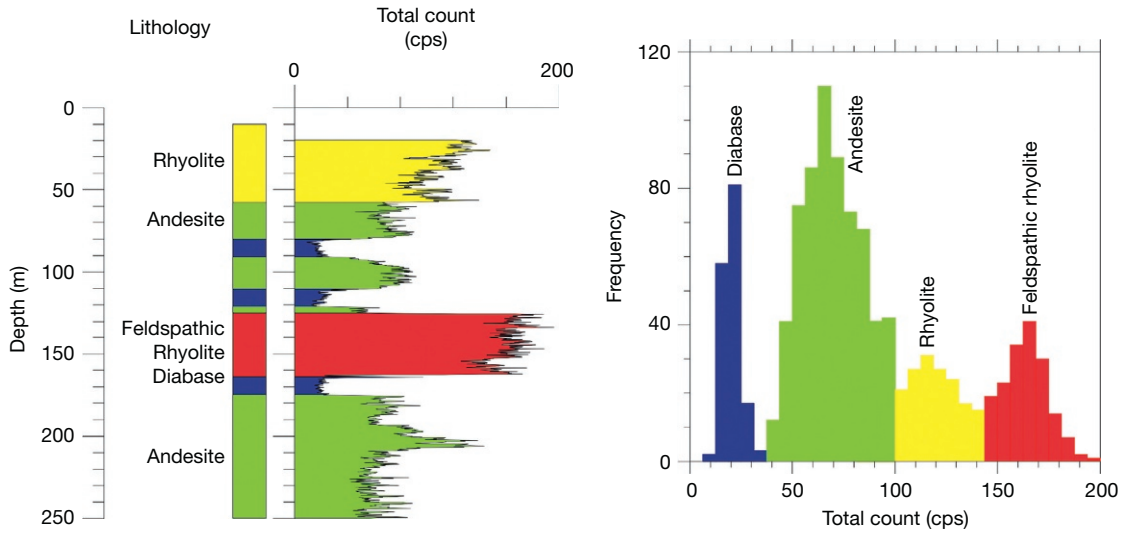


Figure 42 Gamma ray response through a sequence of volcanic rocks, Buchans Mine, Newfoundland. The rhyolite, which contains the most potassic feldspars, shows higher response than the diabase, which has the lowest feldspar content, and the andesite, which has intermediate content.

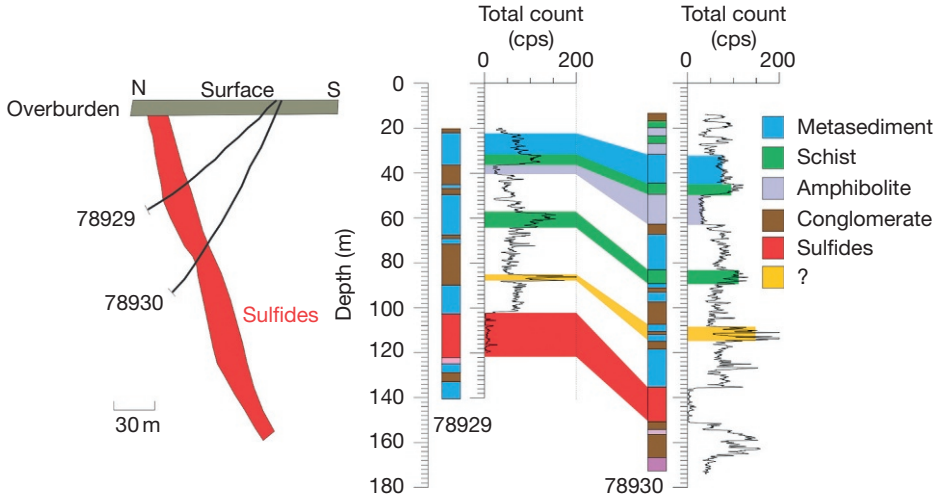


Figure 43 Gamma ray logs acquired in two drill holes at the McConnell nickel deposit, Sudbury, Ontario, Canada, showing their use in stratigraphic correlation. The high gamma ray activity between 110 and 118 m in-hole 78930 (shaded as orange) is identified as metasediments, and a similar response around 90 m in-hole 78929 is in conglomerates.

11.14.8.6.2 Stratigraphic correlation

Often, geologic drill core logging is done by several different geologists in a multihole exploration drilling project. Because of difficulties in visual identification of the different rocks intersected in a borehole, there are often inconsistencies in identifying and classifying different rock types. It has been shown that unraveling ambiguities in lithologic identification can be greatly improved by using natural gamma ray logging. Figure 43 shows an example from a nickel deposit in the Sudbury Basin, Ontario, where TC gamma ray logs were used to unravel and refine the stratigraphy in a metamorphic geologic environment. The gamma ray logs show that geologic units identified as different units in two adjacent drill holes are actually the same. For example, the amphibolites, defined by low gamma ray activity, and the schists by higher gamma ray activity are well correlated between the two drill holes. It is apparent from the gamma ray logs that numerous details are missing from the geologic logs. The gamma ray logs, in this case, provided the geologist with more precise and consistent information to construct a more objective and accurate geologic section at this deposit than was possible from visual examination of the drill core. Spectral logs, K, U, and Th (not shown), provided additional information on the intersected lithology. The radioactivity in the schists, for instance, is primarily from potassium and thorium and virtually no uranium. The high gamma ray activity indicated at about 87 m in drill hole 78929 and around 110 m in drill hole 78930 is primarily uranium enrichment with some thorium, but virtually no potassium. This example shows the possibility of obtaining a detailed radioelement 'fingerprint' of the geology using spectral

gamma ray logs that could provide additional data for more accurate hole-to-hole stratigraphic correlation at the deposit. The use of geophysical logs as an aid in producing geologic logs was described by Killeen et al. (1995).

11.14.8.6.3 Mineral exploration

The applications of natural gamma ray logs in mineral exploration include

1. direct detection of uranium mineralization and ore grade estimation,
2. mapping alteration associated with mineralization (e.g., in gold deposits, porphyry copper deposits, and VMS deposits), and
3. mapping characteristic prospective lithologies (e.g., kimberlite phase and facies classification in diamond exploration).

11.14.8.6.3.1 Direct detection of uranium mineralization

In uranium exploration, natural gamma ray logs are routinely acquired in drill holes to locate uranium mineralization and estimate the uranium grade. The most commonly used log is the TC with the assumption that all of the radioactivity is from uranium and its daughters. This assumption has worked reasonably well in most uranium exploration and delineation projects. However, there are a number of situations where both potassium and thorium contribute significantly to the overall radioactivity that is detected. Pegmatite intrusive deposits generally have significant amounts of thorium (Conaway and Killeen, 1980). Figure 44 shows an example of how using spectral gamma ray logs can significantly improve uranium delineation and ore grade estimation. The TC gamma ray log shows high gamma ray activity in pegmatites. The stripped spectral logs (Ks, Us, and Ths) clearly distinguish pegmatites with uranium mineralization alone (1) from those with a mixture of uranium and thorium (2) and those that are primarily thorium-rich (3). If uranium ore grades were derived only from the TC log, the uranium grade in pegmatite (2) would be overestimated and that of pegmatite (3) would be inaccurate as there is virtually no uranium in this pegmatite.

In the Athabasca unconformity-related uranium deposits of Saskatchewan, Canada, uranium mineralization is often accompanied by both potassium and thorium in both the basement and sandstone-hosted deposits. The TC and spectral gamma ray logs from the McArthur River deposit (Figure 45) show two anomalously high gamma ray activity zones in the Manitou Falls b stratigraphic unit. The spectral logs reveal that the upper anomalous zone on the TC log (330–420 m) is due to an increase in thorium from the presence of the mineral crandallite (Mwenifumbo and Bernius, 2007), while the lower anomalous zone (530–580 m) contains a mixture of uranium and thorium. This example again illustrates the importance of spectral gamma ray logging measurements for accurate in situ uranium grade determinations. The high potassium count rate in the pelites and semipelite in the basement rocks can also affect uranium grade determinations.

11.14.8.6.3.2 Indirect detection of mineralization: mapping alteration

Both TC and spectral gamma ray logs have been effectively used in indirectly detecting mineralization by mapping the associated alteration and thereby increasing the probability of locating deposits. Aluminum-bearing phosphates of the crandallite group, for instance, are reported to be widespread and spatially related to hydrothermal clay paragenesis associated with uranium ore bodies in the Proterozoic basins of Kombolgie in Australia and Athabasca in Canada (Gaboreau et al., 2003; Wilson, 1985). They were also reported to be in hydrothermally altered basement rocks in both places (Madore

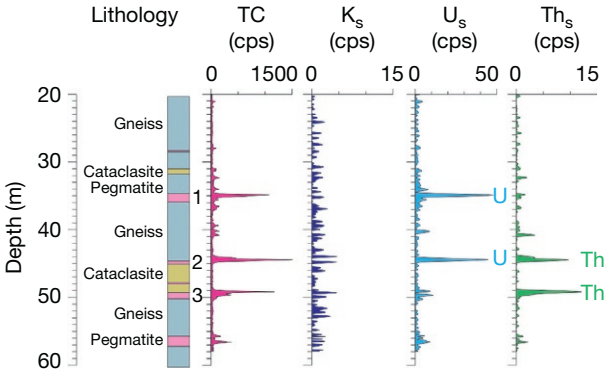


Figure 44 Total count and spectral gamma ray logs acquired at a pegmatite uranium deposit in the Bancroft uranium mining area, Ontario, Canada. K_S, U_S, and Th_S (cps) are stripped count rates in the K, U, and Th windows. The high gamma ray activity in the pegmatite (1) is due to high uranium content. The pegmatite unit (2) has both uranium and thorium. High gamma ray activity in pegmatite (3) is primarily due to thorium. This example shows the utility of spectral gamma ray logs for accurately estimating the uranium content in an environment where both uranium and thorium are present.

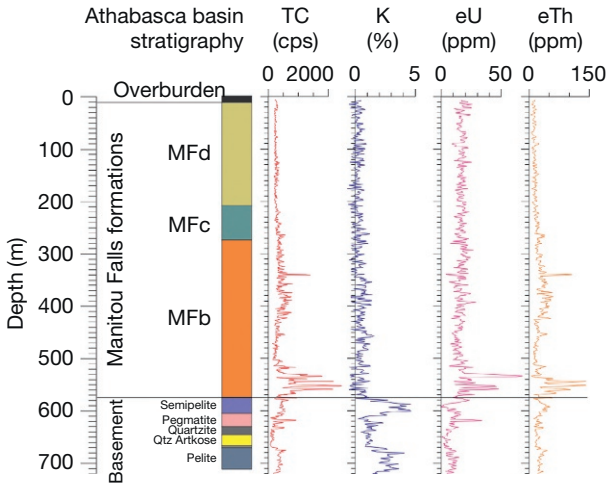


Figure 45 Total count gamma and spectral gamma ray (K, eU, and eTh) logs from the McArthur River uranium deposit in the Athabasca Basin, Saskatchewan. The Athabasca sandstones consist of the Manitou Falls formations (MFd, MFc, and MFb referred to as Manitou Falls d, c, and b Formations, respectively). The gamma ray anomaly above the unconformity between 520 and 580 m (basement–sandstone contact) is due to an increase in both uranium and thorium with a narrow, mainly uranium anomaly around 530 m.

et al., 2000). Potassic alteration, which is also a characteristic of many gold deposits, Cu–porphyry deposits, and VMS deposits, can be delineated by spectral logs. Refer also to **Section 11.14.8.2**.

11.14.8.6.3.3 Crandallite enrichment in the Athabasca Basin

Some anomalously high thorium concentrations in the sandstones were noted in the Athabasca Basin in northern Saskatchewan and Alberta during exploration for unconformity-related uranium deposits. The Th enrichment is found in the conglomeratic strata in the sandstones. The distribution of Th in such coarse-grained fluvial strata is usually associated with accessory minerals in the heavy mineral fraction. Instead, petrographic, chemical, and scanning electron microscope analyses showed that Th in the conglomerate beds resides in a Th-rich aluminophosphate mineral, which belongs to the solid solution series of the crandallite group, including crandallite and goyazite. Crandallite group minerals form in highly acidic environments and are associated with clay minerals such as illite.

The development of these diagenetic and alteration minerals is believed to be associated with hydrothermal alteration processes that operated during the formation of uranium deposits in the Athabasca Basin (Mwenifumbo and Bernius, 2007; Mwenifumbo et al., 2004). **Figure 46** is an example of gamma ray spectral logs, TC, K (%), eU (ppm), and eTh (ppm), and percent grain size >2 mm from drill core logging acquired near the McArthur River uranium deposit in the Athabasca. The elevated eTh content in crandallite-enriched zones is closely correlated with the high percent grain-size log. The sandstones with high percent grain size >2 mm, which are conglomeratic, would have undergone more intense alteration during the mineralization process because of their higher porosity than the normal sandstones.

11.14.8.6.3.4 Potassic alteration in gold deposits

Potassic alteration is one of the most important characteristics of hydrothermal mineral deposits that is useful in exploration. Gold deposits are associated with hydrothermal alteration due to effects of mineralizing fluids. Boyle (1979) gave a comprehensive description of the various types of alteration associated with gold mineralization. The major types of alteration processes reported for most gold deposits include sericitization (potassic), carbonatization, silicification, and pyritization. Sericitization, one of the most common types of wall-rock alteration in gold deposits, consists of the development of potassium mica (generally sericite or hydromuscovite) as a result of the hydration or rearrangement of K, Al, and SiO_2 within intensely altered wall rocks. Chemically, sericitization generally involves the introduction of K or Na and OH or F into the rocks. An increase in K results in a corresponding increase in the radioactive isotope ^{40}K. Zones enriched in potassium form excellent targets for spectral gamma ray exploration methods (Mwenifumbo et al., 1983; Mwenifumbo et al., 2006; Pflug et al., 1996; Urbancic and Mwenifumbo, 1986). Conversely, during the process of carbonatization and silicification, there may be a depletion of potassium (volume reduction) within the altered zones. Hence, carbonatized and silicified zones will exhibit low radioactivity on the gamma ray logs. Refer also to **Section 11.14.8.2**.

Figure 47 shows an example of the use of gamma ray logs in mapping sericitized zones associated with gold mineralization

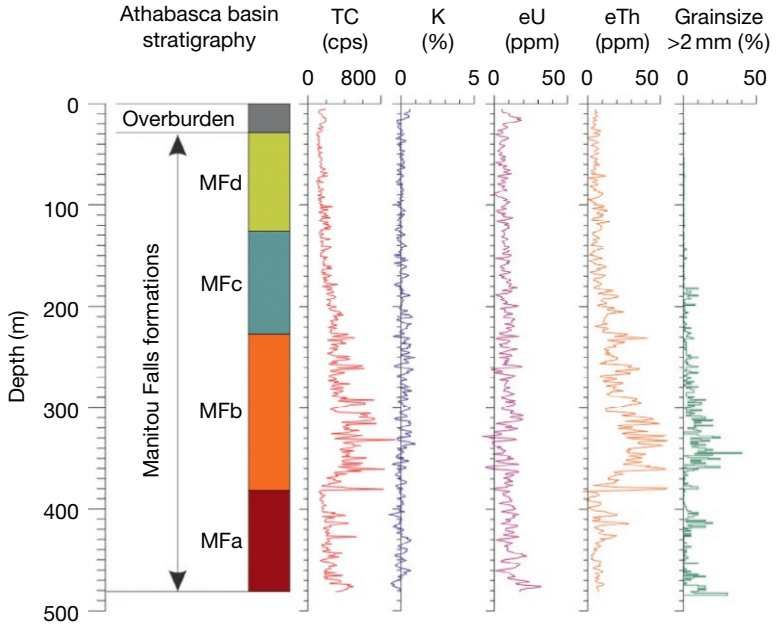


Figure 46 Total count and spectral gamma ray (K, eU, and eTh) logs and percent grain size >2 mm (from the drill core). The high gamma ray activity in the MFb formation is due to the alteration mineral crandallite and correlates with increase in >2 mm grain size, which is a conglomeratic zone. This conglomeratic MFb sandstone (high percent grain size) are more porous and hence subject to increased fluid flow during the alteration process at the time of mineralization. Reproduced from Mwenifumbo CJ and Bernius GR (2007) Crandallite-group minerals: Host of thorium enrichment in the eastern Athabasca Basin, Saskatchewan. In: Jefferson CW and Delaney G (eds.) *EXTECH IV: Geology and Uranium EXploration TECHnology of the Proterozoic Athabasca Basin*, pp. 521–532. Ottawa, ON: Geological Survey of Canada, Bulletin 588.

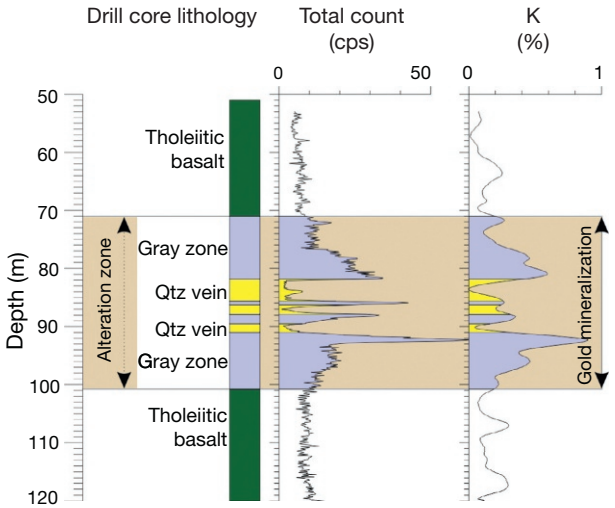


Figure 47 Total count and percent K log across a gold mineralized alteration zone consisting of sericite–chlorite–quartz–carbonate from the Hoyle Pond gold deposit in Timmins, Ontario, Canada. The gray zones are sericite–chlorite alteration zones, which exhibit high total count (high percent K). The low potassium content within the alteration zone outlines the quartz–carbonate veins. The tholeiitic basalts, which are unaltered, exhibit low gamma ray activity.

in the western Abitibi greenstone belt in Ontario, Canada. Gold mineralization is structurally controlled and confined to highly altered shear zones within mafic and massive volcanic flows. Typical alteration within shear zones consists of quartz–carbonate–sulfide (mainly pyrite and arsenopyrite) and sericite–chlorite–carbonate schist.

Similar to the Abitibi belt, the radioelement distribution of K, eU, and eTh is also important in geophysical exploration for gold deposits in the Yellowknife area in the Northwest Territories, Canada. K, eU, and eTh logs can be used to optimize the collection of drill core samples for laboratory gold assaying. In the Yellowknife gold camp, areas with high concentrations of both K and Th are associated with changes in lithology, not alteration. However, areas showing high K concentration alone are associated with alteration and gold mineralization.

11.14.8.6.4 Kimberlite characteristics and classification: diamond exploration

Several studies on the abundance of uranium and thorium in kimberlite have indicated that these radioelements are enriched compared to other ultramafic and mafic rocks (Ahrens et al., 1967; Kresten, 1974; Lutts and Mineeva, 1973; Paul et al., 1977). The U and Th concentrations and Th/U ratios have been used to differentiate kimberlites from ultramafic and basic rocks. The U and Th content also varies considerably within individual pipes and among the different pipes, primarily due to the different facies (crater, diatreme, and hypabyssal) and source material of the eruptive phases of kimberlitic intrusions (Mwenifumbo and Kjarsgaard, 2000).

It has also been suggested that U and Th and their ratios may be used to distinguish between barren and diamondiferous kimberlites (Lutts and Mineeva, 1973). These radioelements may also provide information on the evolution of kimberlites, the type of crustal contamination, and the alteration processes during and after emplacement (Fesq et al., 1975; Kresten, 1974; Paul et al., 1977). The possibility of using GRS to prospect for and delineate kimberlite was first suggested by Lutts and Mineeva (1973). However, there are little data published on the subject. There have been a few reported case studies in Australia describing airborne and ground spectral gamma ray surveys (Atkinson, 1989; Macnae, 1979). At present, there are very few published data on the potassium, uranium, and thorium distributions in Canadian kimberlites. In Canada, Mwenifumbo and Kjarsgaard (2000) presented the first case study on the distribution of K, U, and Th within a single kimberlite pipe in the Fort à la Corne kimberlite field in Saskatchewan. Several eruptive phases that were established from geologic, geophysical, and geochemical observations were correlated to the variations in gamma ray activity observed on the gamma ray logs. The crater facies kimberlites (eruptive pyroclastics) in the Fort à la Corne kimberlite (see **Figure 31**) show variations in the TC gamma ray activity that distinguishes different eruptive phases. These variations suggest that there were different sources of kimberlitic material during its emplacement history. Spectral gamma ray logs and geochemical data indicate that the gamma ray activity is mainly due to variations in the Th content of these rocks (Mwenifumbo and Kjarsgaard, 2000; Mwenifumbo et al., 1996a).

Figure 48 shows another example of the use of gamma ray logs to map different facies in a kimberlite pipe in the Kirkland Lake area of Ontario, Canada. The gamma ray log clearly improves the delineation of the kimberlite pipe intersection. Four distinct gamma ray signatures are identified and delineated within the lithic tuffisitic kimberlite breccia (LTKB) between 48 and 204 m: (1) LTKB, (2) silicified LTKB, (3) tuffisitic kimberlite breccia (TKB), and (4) kimberlite breccia. The thin, silicified LTKB between 92.8 and 94.2 m has an abnormally high gamma ray activity compared to the rest of the LTKB unit. The increase in gamma ray activity is probably due to enrichment in uranium along a fracture zone (West and Laughlin, 1976). Generally, gamma ray activity is lower in the TKB and higher in kimberlite, which is composed of massive igneous rock that is less brecciated and less porous than the TKB zones.

11.14.8.7 Environmental Radiation Monitoring

While GRS was initially developed as a uranium exploration technique, its applications have evolved to include bedrock and surficial geologic mapping and direct or indirect support for mineral exploration for a variety of mineral commodities. In addition, GRS has also been effectively used to support a number of environmental radiation studies and monitoring activities. The IAEA (1995, 2003) provides a review of a number of these applications including the environmental monitoring of nuclear power plants, the mapping of areas contaminated by nuclear accidents, the search for lost radioactive sources, the mapping of the concentration and redistribution of natural radiation material due to the extraction and processing of uranium mining-related activities, emergency radiation response, and security-related activities. Similarly, regional AGRS survey data acquired to support mapping and mineral resource assessments have also been used to support radon potential mapping and risk assessment activities.

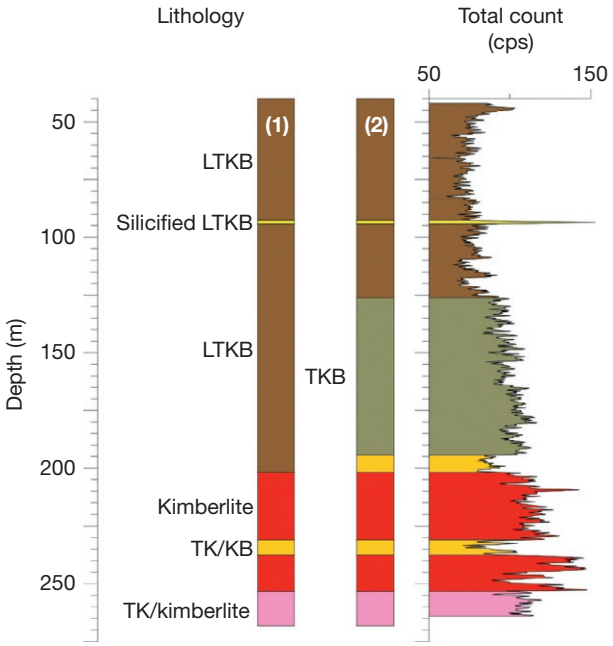


Figure 48 Gamma ray log from the Kirkland Lake kimberlite cluster in Ontario, Canada. Lithology in column (1) is from drill core logging and column (2) is an interpretation based on drill core and gamma ray log analysis. The gamma ray log is able to refine the lithology in a lot more detail within the lithic tuffisitic kimberlite breccia (LTKB) unit between 48 and 204 m. This LTKB unit can be subdivided into three units; LTKB, TKB, and TK/KB with the gamma ray activity being highest in the TKB unit. The kimberlite unit, which is composed of massive igneous rock that is less brecciated and less porous than the tuffisitic kimberlite breccia zones (hypabyssal kimberlite facies), has the highest gamma ray activity due to its intrinsic high radioelement concentration. KB, kimberlite breccia; TK, tuffisitic kimberlite; TKB, tuffisitic kimberlite breccia. Reproduced from Mwenifumbo CJ, Killeen PG, and Elliott BE (1998) Borehole geophysical signatures of kimberlites in Canada. *The Log Analyst* 39: 38–52.

11.14.8.7.1 Radon potential mapping

Radon (^{222}Rn) is a naturally occurring, colorless, odorless, and tasteless radioactive gas. It is a daughter product in the ^{238}U decay series and has a half-life of 3.82 days, sufficiently long to migrate from rocks and soils into buildings. Soil gas radon originating from rocks and soils and contributing to the indoor environment is the main source of natural radiation exposure to the population and is the second leading cause of lung cancer (World Health Organization, 2009). Indoor radon levels can only be determined by a direct measurement of the indoor air. Commonly, this is done using an alpha track radon detector placed in a suitable living area of the home for a long-term sampling period of between 60 and 90 days. Baseline geoscience data, including AGRS, bedrock, and surficial geology, are often used, in conjunction with indoor radon measurements, as primary datasets to identify areas prone to elevated levels of indoor radon. AGRS has been widely used to indirectly assess the radon potential in many countries including, but not limited to, the United States (Otton et al., 1995; White et al., 1993), the United Kingdom (Ball et al., 1995), Sweden (Akerblom, 1995), Norway (Smethurst et al., 2008), Czech Republic (Barnet, 1995; Barnet et al., 2008), and Canada (Cocksedge et al., 1993). Most AGRS surveys include a measurement of atmospheric radon levels that is required for correction of the U channel data (see **Sections 11.14.5.3.4–11.14.5.3.6**). However, these data generally have insufficient resolution to be used to identify radon-prone areas.

In the Oslo region of Norway, Smethurst et al. (2008) found a clear spatial correlation between areas with elevated AGRS eU signatures and areas with a higher percentage of houses with average indoor radon concentrations above the established radon action level of 200 Bq m^{-3}. They concluded that AGRS is a primary dataset for radon risk assessment and the most effective dataset in areas with few or no indoor measurements. In Canada, Cocksedge et al. (1993) successfully used a variety of geoscience datasets including AGRS, lake and stream sediment geochemistry, and geology to predict which native communities across Canada were likely to have higher indoor radon concentrations.

At Oka, Quebec, Ford et al. (2001) found that increasing eU concentrations, measured by an AGRS survey, showed a progressive and clear association with increasing indoor radon concentrations and an increasing percentage of homes that exceeded established indoor radon levels (**Figure 49**; **Table 7**). These results were used by public health officials to identify specific areas at risk of overexposure to indoor radon and supported a number of other land use and public health initiatives.

While direct indoor testing for radon is the only method to identify individual dwellings that exceed established indoor radon guidelines, the association between high eU concentrations measured by AGRS and high indoor radon concentrations demonstrates the effectiveness of AGRS as a predictive tool for the identification of radon-prone areas, including areas with no residential development.

11.14.8.7.2 Mapping fallout from nuclear incidents

Several AGRS data analysis techniques have been developed to detect and map man-made sources of radioactivity dispersed from nuclear bomb tests, nuclear power plants, and nuclear-powered satellites. The most important of these man-made radioactive sources are listed in **Table 8**. An important fact about these radionuclides that is pertinent for AGRS is that most emit gamma rays with energies below those associated with the common, naturally occurring radioelements, K, eU, and eTh at 1461, 1765, and 2615 keV, respectively.

On 24 January 1978, nuclear-powered satellite Kosmos 954 re-entered the Earth's atmosphere and scattered radioactive debris over a large area of northern Canada. **Figure 50** shows the distribution of total radioactivity measured by AGRS along a series of profiles flown across the frozen surface of Great Slave Lake in northern Canada several weeks after the reentry of the satellite. Portions of the profiles north of the satellite's wind-corrected trajectory show the typically low overwater background levels. Those portions south of the trajectory show the diminishing downwind dispersion of radioactive debris. The general search area is characterized by high levels of natural radioactivity, a portion of which is also shown in **Figure 50**. Bristow (1978) described how AGRS was instrumental in the discovery of radioactive debris from the satellite. A ratio of low-energy (300–1400 keV) gamma ray count rate to high-energy (1400–2800 keV) gamma ray count rate was found to remain

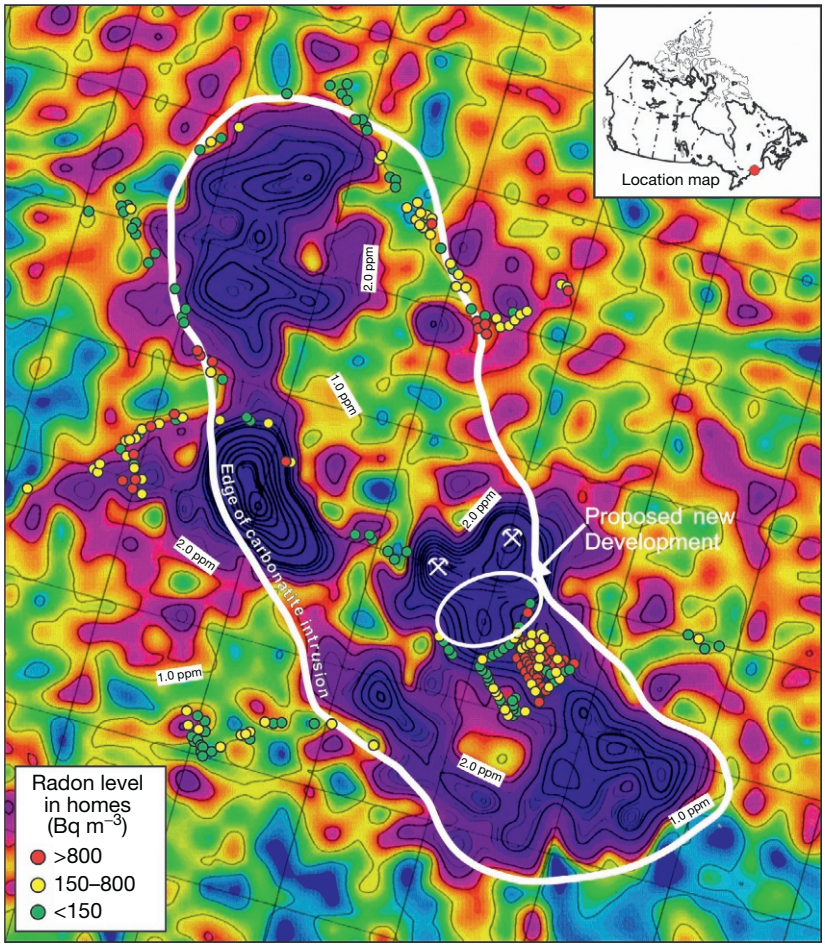


Figure 49 Equivalent uranium (ppm) map from a 200 m line spaced AGRS survey of the Oka, Quebec area. Edge of carbonatite intrusion outlined by white line. Colored dots show indoor radon measurements used to establish association with AGRS eU (ppm) values. Scale provided by 1 km UTM grid. Reproduced from Ford KL, Savard M, Dessau JC, Pellerin E, Charbonneau BW, and Shives RBK (2001) The role of gamma ray spectrometry in radon risk evaluation: A case history from Oka, Quebec. *Geoscience Canada* 28: 59–64.

Table 7 Percentage of homes exceeding established Canadian and US indoor radon levels classified by airborne eU concentrations

AGRS eU (ppm)	*<150 Bq m^{-3}*	*>150 Bq m^{-3}[a]*	*>800 Bq m^{-3}[b]*
0–1(35)	54.3% (19)	45.7% (16)	0% (0)
1–2(78)	48.7% (38)	51.3% (33)	9.0% (7)
2–4(42)	52.4% (22)	47.6% (11)	21.4% (9)
4–6(45)	24.4% (11)	75.6% (14)	44.4% (20)
>6(42)	19.1% (8)	80.9% (15)	45.2% (19)

[a]United States Environmental Protection Agency indoor radon guideline.
[b]Health Canada 2001 indoor radon guideline, lowered to 200 Bq m^{-3} in 2007.
Parentheses indicate number of homes in each category.
Source: Ford KL, Savard M, Dessau JC, Pellerin E, Charbonneau BW, and Shives RBK (2001) The role of gamma ray spectrometry in radon risk evaluation: A case history from Oka, Quebec. *Geoscience Canada* 28: 59–64.

Table 8 Principal man-made nuclides

Isotope	*Half-life (days)*	*Energy of gamma rays (keV)*
^{95}Z	65	724, 756
^{95}Nb	35	765
^{99}Mo	3	740
^{103}Ru	40	497
^{106}Ru	368	512
^{131}I	8	364
^{132}Te	3	230
^{134}Cs	730	605, 795
^{137}Cs	11 000	662
^{140}Ba/La	13	1596

Source: Grasty RL (1995) Environmental monitoring by airborne gamma ray spectrometry, experience at the Geological Survey of Canada. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 93–102. TECDOC-827. Vienna: IAEA.

constant for areas of natural radioactivity and to increase markedly over areas with significant low gamma ray energy debris from the satellite (Grasty, 1995). This technique permitted the man-made radioactive debris to be successfully located by AGRS. Subsequent analysis of AGRS spectra collected over the frozen surface of Great Slave Lake confirmed the detection of gamma rays characteristic of nuclear fission products. Figure 51(a) (Grasty, 1995) shows a standard natural gamma ray spectrum; Figure 51(b) shows the residual spectra following removal of the natural component.

Analysis of AGRS data was also effective for mapping radioactive fallout from the Chernobyl nuclear reactor accident.

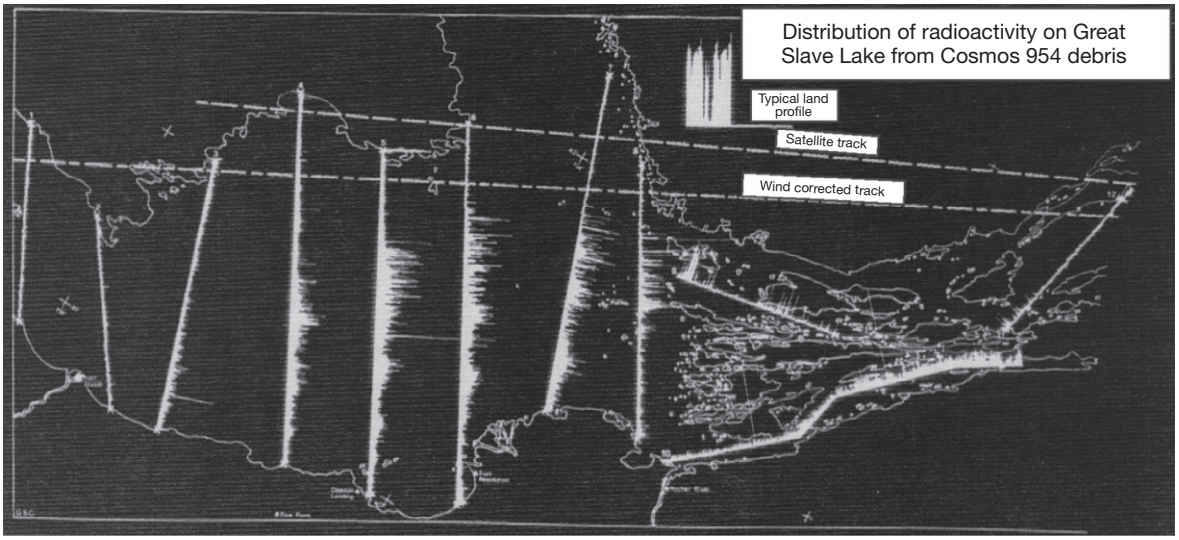

Distribution of radioactivity on Great Slave Lake from Cosmos 954 debris
Typical land profile
Satellite track
Wind corrected track

Figure 50 Airborne gamma ray spectrometric total count profiles collected by the Geological Survey of Canada over the frozen surface of Great Slave Lake, Northwest Territories. Radioactive debris from Kosmos 954 shows a diminishing downwind distribution from the wind-corrected trajectory. Wind direction was from north to south. Normal overwater background signature visible over portions of flight lines that extend north of the trajectory. Reproduced from Shives RBK, Ford KL, and Charbonneau BW (1995) Geological Survey of Canada Workshop Manual: Applications of gamma ray spectrometric/magnetic/VLF-EM surveys, p. 85. Open File 3061. Ottawa, ON: Geological Survey of Canada.

Grasty (1995) described the technique to separate predominantly low-energy man-made gamma ray signatures from higher-energy, natural gamma radiation from K, eU, and eTh sources by determining the appropriate stripping ratios required to remove the naturally occurring gamma ray contributions to the lower energy channels. Similarly, Sanderson et al. (1995b) and Sanderson et al. (1995a) described how many of the standard calibration methods for determining stripping ratios, attenuation coefficients, and sensitivity corrections are extendable to other man-made radionuclides.

11.14.8.7.3 Monitoring nuclear power facilities

Hetu et al. (1993) presented results from a 1990 AGRS survey in the vicinity of a nuclear power station near Kozloduy, Bulgaria. Analysis of these data showed elevated levels of ^{137}Cs and ^{60}Co in soils adjacent to cooling discharge canals from the power station. The appropriate coefficients to remove (strip) the natural contributions from K, eU, and eTh to the ^{137}Cs and ^{60}Co windows were determined and applied to the data. These data were then processed according to established procedures (IAEA, 1991) to produce maps for the standard measured and derived AGRS products together with maps that showed the distribution of ^{137}Cs and ^{60}Co. **Figure 52** shows the distribution of ^{137}Cs adjacent to the cooling discharge canals.

Grasty and Sander (2001) presented results from a detailed, helicopter AGRS survey over the Darlington and Pickering nuclear power facilities in Canada on the shore of Lake Ontario. Survey results documented the background radiation levels and radiation signatures at and around the nuclear power facilities. In addition to producing maps for the naturally occurring radioelements (K, eU, and eTh), the AGRS data were processed to determine levels of ^{137}Cs, ^{60}Co, ^{16}N, and ^{41}Ar. Survey results provide baseline information that could become important should an accident occur and are useful for community consultation purposes. AGRS systems have also been employed to monitor and map, in near real time, active low-level emissions from nuclear power facilities. Shives et al. (1995) presented results from an AGRS survey flown near the Douglas Point nuclear power facility. **Figure 53** shows an exposure rate map over Lake Huron, Ontario, near the Douglas Point facility. The distribution of ^{41}Ar in the air, downwind from the generating station, is clearly shown. The maximum, calculated exposure rate at ground level was found to be about three times the average natural background radiation in Canada, but no higher than levels due to naturally occurring radioelements (K, eU, and eTh) found in many parts of Canada (Grasty, 1983; Grasty et al., 1984).

Stationary GRS monitoring systems are often employed around nuclear power facilities and other industrial sites including waste disposal and scrap steel recycling sites to provide real-time monitoring. Cox (1995) provided a number of examples where radioactive sources were accidently processed with recycled scrap steel resulting in contamination of the new products. Often, the contamination was not discovered until months or years later and required significant costs to decontaminate.

11.14.8.7.4 Emergency radiation response and preparedness

Many countries, in response to nuclear accidents like Chernobyl and Fukoshima and/or the possibility of intentional radiological dispersions such as a dirty bomb, have established emergency radiation response programs. These monitor not only background radiation due to natural radiation sources but also existing man-made sources from previous nuclear testing and accidental dispersions. National field exercises, to test and modify emergency response capabilities and readiness in the event of a nuclear or radiological incident, are held regularly in some countries. Such exercises are also important to test the interoperability of different government agencies that often make up such emergency response programs. Internationally, a number of joint exercises have been held to

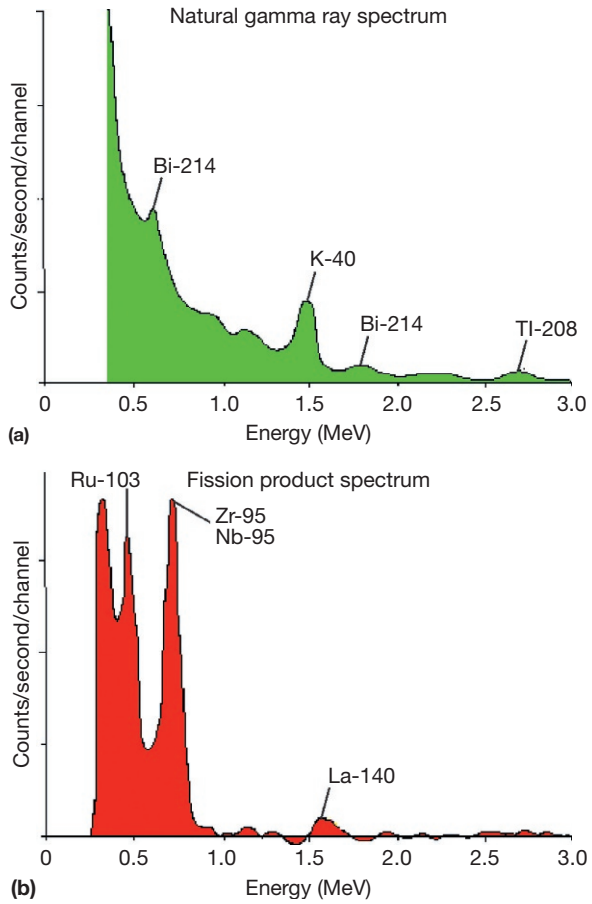


Figure 51 Standard natural gamma ray spectrum (a) compared with a fission product spectrum (b). Fission product spectrum collected by the Geological Survey of Canada AGRS system from radioactive debris scattered on the ice of Great Slave Lake (From http://www.collectionscanada.gc.ca/webarchives/20061103073032 and http://gsc.nrcan.gc.ca/gamma/index_e.php).

further test and improve methods for mapping radiological contamination. In 1995, the Nordic Nuclear Safety Research Secretariat (NKS) held a multicountry exercise, RESUME 95, to test the effectiveness of a number of in situ, car-borne, and airborne systems to map radiological contamination levels and to locate lost radioactive sources (NKS, 1997). Similarly, Sanderson and Ferguson (1997) described results from a European Community-sponsored project that focused on the effectiveness of AGRS systems, from several countries, to monitor and map radiological fallout. Such exercises have clearly demonstrated the central role that AGRS plays in emergency radiation response programs.

There are now a number of commercially available AGRS systems that are particularly suited for emergency radiation mapping. System capabilities often include integrated GPS, laser, and/or radar altimeter along with full multichannel spectral recording for individual detectors. The ability to interface auxiliary detectors such as neutron or other radiation detectors is important as is the ease of installation into a variety of survey platforms and the capability to record all radiation and ancillary data during acquisition for subsequent playback and real-time telemetry data, including spectra and ancillary data. Monitoring software that includes real-time and post-acquisition display of radiation and positional data in a variety of user-selectable formats, at the point of collection and, remotely, off-site, is desirable as is real-time and post-acquisition radionuclide identification, both in vehicle and, remotely, for nuclear, medical, and industrial sources, and the ability to export data in a variety of formats for import into standard geophysical and GIS packages.

11.14.9 Conclusions and Future Directions

Radiometric techniques have evolved into valuable near-surface geophysical methods that can be applied from

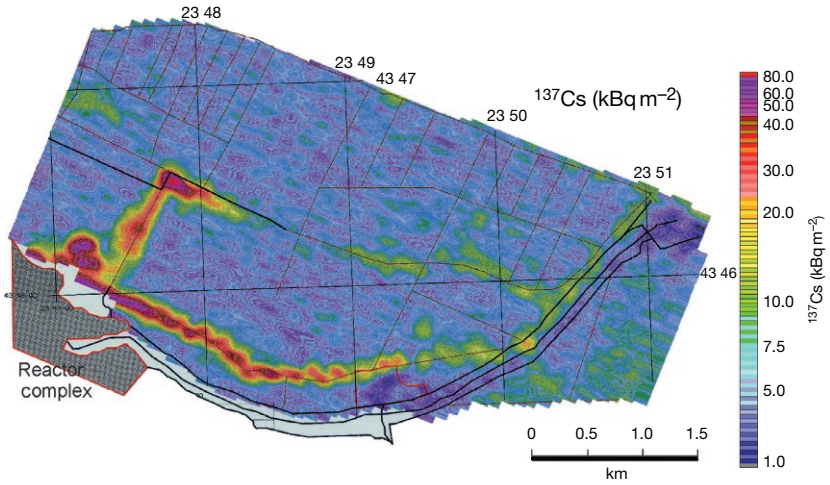


Figure 52 A map of the Kozloduy area, Bulgaria, showing deposition of ^{137}Cs in the vicinity of a nuclear reactor. Reproduced from Hetu R, Rangelov R, Grant JA, Grasty RL, and Graham D (1993) Monitoring natural and man-made radiation by airborne gamma ray spectrometry, Kozlodui Area, Bulgaria, 1990. Open File 2573, Geological Survey of Canada.

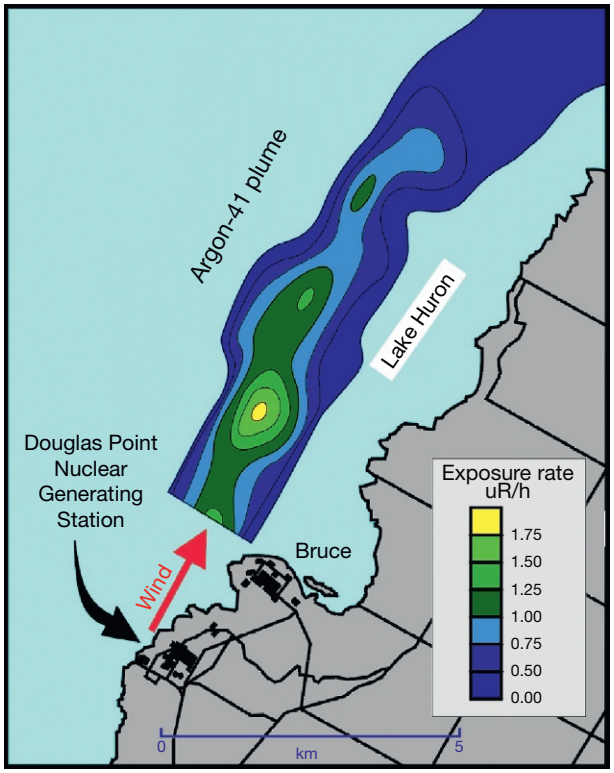


Figure 53 Airborne gamma ray spectrometric map of the area around the Douglas Point Nuclear Generating Station showing the distribution of ^{41}Ar in the air over Lake Huron downwind from the generating station (From http://www.collectionscanada.gc.ca/webarchives/20061103073032 and http://gsc.nrcan.gc.ca/gamma/index_e.php).

airborne, surface, and borehole platforms. The field of radiometrics is dominated by the passive technique of GRS used for quantitative measurement of natural radioactivity. Although some 'active' methods (which use radioactive sources) are in use on the ground surface (e.g., XRF) and in boreholes (e.g., gamma–gamma density logging), the methods are not as broadly applicable to geologic and environmental investigations as GRS.

Since the 1960s, data acquisition hardware for GRS has developed into sophisticated systems designed to extract maximum information from the measured natural gamma ray spectra of K, U, and Th, minimizing noise and maximizing spatial resolution (e.g., in airborne surveys). Methods and facilities for accurate calibration of spectrometers are available for airborne, surface, and BGRS. The parameters that affect the measurement results (e.g., radon background, cosmic radiation, air and water attenuation, temperature, pressure, and counting statistics) are well understood, and appropriate data corrections have been developed.

Knowledge of the relationship between the natural radioelement distribution and geologic/environmental processes has reached an advanced level. GRS has been applied in geology to bedrock and surficial geologic mapping and mineral exploration for a variety of deposit types including U, Th, gold, massive sulfides, diamond in kimberlite, porphyry, and polymetallic IOCG-type deposits. It has been applied to environmental investigations such as mapping areas prone to high indoor radon concentrations, mapping radioactive fallout from man-made sources, and monitoring nuclear power facilities. Innovative applications continue to be developed that include, for example, the mapping of soils to measure the amount of fertilizer (from the K content) being applied on farm fields to develop regulations to limit pollution of adjacent rivers and streams.

A number of data presentation and visualization methods have been used to enhance the interpretation of gamma ray spectrometric data. Multiple parameters have also been combined, especially radar methods, to add to the uniqueness of interpretations by imposing constraints on multiple possibilities. When available, AGRS data have become a standard input dataset for RPM exercises performed in advance of field geologic mapping and mineral exploration activities.

Future advances in GRS will likely be in the area of data processing and data presentation and may include the more general use of the full spectral data to continue to extract maximum information by minimizing noise, along with the possible development of corrections for terrain effect (i.e., surface geometry). Additional enhancement of data presentations and merging and correlation with other geophysical parameters for 3-D modeling are possible new directions. Radiometric data may be used to constrain interpretations in some cases. As the number of government-sponsored regional surveys and industry-sponsored targeted exploration-oriented surveys acquires more GRS data, it is likely that the number and types of applications will grow. The hardware associated with GRS has seen numerous major developments along with the evolution of electronic and computing power to where it is today. Some advances in data acquisition technology may arise from the development of new detectors with new materials and possibly in new placement configurations of detectors in the case of airborne spectrometry to reduce internal shielding between them. Another potential new direction for AGRS methods will be to exploit the differential sensitivities of arrays of detectors caused by the internal shielding. These include using the relatively new technology involved in directional arrays and gamma camera devices to develop systems with some form of directionality (e.g., a horizontal cross flight line radiometric gradient).

Acknowledgments

The three authors, who have over 85 years of combined experience working for the Geological Survey of Canada, would like to acknowledge the wide range of expertise of members of the GSC's Radiometric, Borehole Geophysics, and Instrumentation Development teams. These teams significantly advanced the development of gamma ray spectrometry and its use in field applications since the 1960s. We wish to express our gratitude to them collectively, for their work, which forms the foundation of this chapter. In particular, we wish to thank John M. Carson of the GSC who provided support and constructive advice throughout its preparation and many revisions. The manuscript has also benefitted from critical reviews by Martin Bates of Sander Geophysics Ltd. and Robert B.K. Shives of GamX Inc.

References

Adams JAS and Clark RB (1972) Computer modeling and experimental calibration of airborne gamma spectrometer systems. In: Adams JAS, Lowder WM, and Gesell TF (eds.) *The Natural Radiation Environment II*, pp. 641–648. Springfield, VA: US Department of Commerce.

Adams JAS and Fryer GE (1964) Portable gamma ray spectrometer for field determination of thorium, uranium and potassium. In: Adams JAS and Lowder WM (eds.) *The Natural Radiation Environment*, pp. 577–596. Chicago, IL: University of Chicago Press.

Adams JAS and Gasparini P (1970) *Gamma Ray Spectrometry of Rocks*. Amsterdam: Elsevier Publishing Company.

Adams JAS and Lowder WM (eds.) (1964) *The Natural Radiation Environment*, p 1069. Chicago, IL: University of Chicago Press.

Adams JAS and Weaver CE (1958) Thorium-to-uranium ratios as indicators of sedimentary processes: Example of concept of geochemical facies. *Bulletin of the American Association of Petroleum Geologists* 42: 387–430.

Ahrens LH, Cherry RD, and Erlank AJ (1967) Observations on the Th–U relationships in zircon from granitic rocks and from kimberlites. *Geochimica et Cosmochimica Acta* 31: 2379–2387.

Akerblom G (1995) The use of airborne radiometric and exploration survey data and techniques in radon risk mapping in Sweden. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 159–180. Vienna: International Atomic Energy Agency, TECDOC-827.

Atkinson WJ (1989) Diamond exploration philosophy, practice, and promises. In: Ross J (ed.) *Kimberlites and Related Rocks*, pp. 1075–1107. Melbourne: Blackwell.

Austin, S.R. (1975). A laboratory study of radon emanation from domestic uranium ores. In: *Radon in Uranium Mining; Panel Proceedings*, Washington, DC, 1973, p. 151. Vienna: IAEA.

Ball TK, Cameron DG, Colman TB, and Roberts PD (1995) The use of uranium exploration data for mapping radon potential in the UK – Advantages and pitfalls. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 139–150. Vienna: International Atomic Energy Agency, TECDOC-827.

Barnet I (1995) The role of airborne gamma spectrometric data in the radon programme of the Czech Republic. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 151–158. Vienna: International Atomic Energy Agency, TECDOC-827.

Barnet I, Pacherova P, Neznal M, and Neznal M (2008) In: *Radon in Geological Environment – Czech Experience*, p 70. Prague: Czech Geological Survey, Special Paper 19.

Barretto PMC (1975) Radon-222 emanation characteristics of rocks and minerals. In: *Radon in Uranium Mining, Panel Proceedings*, Washington, 1973, p. 129. Vienna: IAEA.

Belknap WB, Dewan JT, Kirkpatrick CV, Matt WE, Pearson AJ, and Robson WR (1959) API calibration facility for nuclear logs. In: *American Petroleum Institute Proceedings*, pp. 289–317.

Blusson SL (1971) *Sekwi Mountain Map Area, Yukon Territory and District of MacKenzie (NTS 105P)*. p. 17. Ottawa, ON: Geological Survey of Canada, Paper 71-22.

Bowie SHU, Hale FH, Ostle D, and Beer KE (1955) Radiometric surveying with a carborne counter. *Bulletin of the Geological Survey of Great Britain* 10: 1–23.

Boyle RW (1979) The geochemistry of gold and its deposits. *Geological Survey of Canada Bulletin* 280: 584.

Bristow Q (1978) The application of airborne gamma ray spectrometry in the search for radioactive debris from the Russian satellite COSMOS 954 (Operation "Morning Light"). In: *Current Research, Part B*, pp. 151–162. Ottawa, ON: Geological Survey of Canada, Paper 78-1B.

Bristow Q (1979) Gamma ray spectrometric methods in uranium exploration: Airborne instrumentation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 135–146. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

Bristow Q, Conaway JG, and Killeen PG (1984) Application of inverse filtering to gamma ray logs: A case study. *Geophysics* 49: 1369–1373.

Broome J, Carson JM, Grant JA, and Ford KL (1987) *A Modified Ternary Radioelement Mapping Technique and Its Application to the South Coast of Newfoundland*. Ottawa, ON: Geological Survey of Canada, Paper 87-14, scale 1:500,000.

Buckle JL, Carson JM, Miles WF, Ford KL, Fortin R, and Delaney G (2011) Geophysical series, geophysical compilation northern Saskatchewan; Geological Survey of Canada, Open File 7039; Saskatchewan Ministry of Energy and Resources, Open File 2011-58; scale 1:750,000.

Campbell JE, Klassen RA, and Shives RBK (2007) Integrated field investigations of airborne radiometric data and drift composition, Nuclear Energy Agency-International Atomic Energy Agency Athabasca test area, Saskatchewan. In: Jefferson CW and Delaney G (eds.) *EXTECH IV: Geology and Uranium Exploration Technology of the Proterozoic Athabasca Basin, Saskatchewan and Alberta*, pp. 533–554. Ottawa, ON: Geological Survey of Canada, Bulletin 588.

Carson JM, Holman PB, Ford KL, Grant JA, and Shives RBK (2003) Airborne gamma ray spectrometry compilation, Nova Scotia. Geological Survey of Canada, Open Files 4464-4471, scale 1:500 000.

Charbonneau BW (1988) Gamma spectrometric and magnetic anomalies associated with Cu–U mineralization, Faber Lake volcanic belt, District of MacKenzie, Northwest Territories. In: *Current Research, Part C*, pp. 255–258. Ottawa, ON: Geological Survey of Canada, Paper 88-1C.

Charbonneau BW and Darnley AG (1970) A test strip for calibration of airborne gamma ray spectrometers. In: *Report of Activities, Part B*, pp. 27–32. Ottawa, ON: Geological Survey of Canada, Paper 70-1B.

Charbonneau BW and Ford KL (1979) Discovery of two uranium occurrences in paleozoic sedimentary rocks at South March, Ontario and South Maitland, Nova Scotia by airborne gamma ray spectrometry and comparison with surveys over Precambrian terrain. *Journal Canadian Society of Exploration Geophysicists* 15.

Charbonneau BW, Killeen PG, Carson JM, Cameron GW, and Richardson KA (1976) The significance of radioelement concentration measurements made by airborne gamma ray spectrometry over the Canadian Shield. In: *Proceedings of International Symposium on Exploration for Uranium ore Deposits. Proceeding Series*, pp. 35–54. Vienna: IAEA.

Charbonneau BW and Legault MI (1994) Interpretation of airborne geophysical data for the Thor Lake area, Northwest Territories. In: Sinclair WD and Richardson DG (eds.) *Studies of Rare-Metal Deposits in the Northwest Territories*, pp. 17–31. Ottawa, ON: Geological Survey of Canada, Bulletin 475.

Charbonneau BW and Swettenham SS (1986) Gold occurrence in radioactive calc-silicate float at Sandybeach Lake, Nueltin Lake area, District of Keewatin. In: *Current Research, Part A*, pp. 803–808. Ottawa, ON: Geological Survey of Canada Paper 86-1A.

Chase GD and Rabinowitz JL (1968) *Principles of Radioisotope Methodology*, p. 633 Minneapolis, MN: Burgess Publishing Company.

Clark RB, Duval JS, and Adams JAS (1972) Computer simulation of an airborne gamma ray spectrometer. *Journal of Geophysical Research* 77: 3021–3031.

Clark SP Jr, Peterman ZE, and Heier KS (1966) Abundance of uranium, thorium and potassium. In: *Handbook of Physical Constants*, pp. 521–541. Boulder, CO: Geological Society of America, Memoir 97.

Cocksedge W, Rankin W, Tostowaryk K, Charbonneau BW, and Grasty RL (1993) National native home radon survey – Maximizing resources through radon potential assessment. In: *Proceedings of the 26th Midyear Topical Meeting of the Health Physics Society, Environmental Health Physics*, Lake Coeur d'Alene, Idaho, 24–28 January 1993, pp. 391–402.

Conaway JG (1978) Problems in gamma ray logging: Thin zone correction factors. In: *Current Research, Part C*, pp. 19–21. Ottawa, ON: Geological Survey of Canada, Paper 78-1C.

Conaway JG (1979a) Problems in gamma ray logging: The effect of dipping beds on the accuracy of ore grade determinations. In: *Current Research, Part A*, pp. 41–44. Ottawa, ON: Geological Survey of Canada, Paper 79-1A.

Conaway JG (1979b) Computer processing of gamma ray logs: A program for the determination of radioelement concentrations. In: *Current Research, Part B*, pp. 27–32. Ottawa, ON: Geological Survey of Canada, Paper 79-1B.

Conaway JG (1980a) Exact inverse filters for the deconvolution of gamma ray logs. *Geoexploration* 18: 1–14.

Conaway JG (1980b) Direct determination of the gamma ray logging system response function in field boreholes. *Geoexploration* 18: 187–199.

Conaway JG (1980c) Uranium concentrations and the system response function in gamma ray logging. In: *Current Research, Part A*, pp. 77–87. Ottawa, ON: Geological Survey of Canada, Paper 80-1A.

Conaway JG (1981) Deconvolution of gamma ray logs in the case of dipping radioactive zones. *Geophysics* 46: 198–202.

Conaway JG (1986) Spatial deconvolution applied to gamma ray logging: A review. In: Killeen PG (ed.) *Borehole Geophysics for Mining and Geotechnical Applications*, pp. 357–360. Ottawa, ON: Geological Survey of Canada, Paper 85-27.

Conaway JG, Bristow Q, and Killeen PG (1980) Optimization of gamma ray logging techniques for uranium. *Geophysics* 45: 292–311.

Conaway JG and Killeen PG (1978a) Quantitative uranium determinations from gamma ray logs by application of digital time series analysis. *Geophysics* 43: 1204–1221.

Conaway JG and Killeen PG (1978b) Computer processing of gamma ray logs: Iteration and inverse filtering. In: *Current Research, Part B*, pp. 83–88. Ottawa, ON: Geological Survey of Canada, Paper 78-IB.

Conaway JG and Killeen PG (1980) Gamma ray spectral logging for uranium. *Canadian Mining and Metallurgical Bulletin* 73: 115–123.

Corriveau L (2007) Iron oxide copper-gold deposits: A Canadian perspective. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration*

Methods, pp. 307–328. St. John's, NL: Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

Cox JR (1995) Radiation detection of buried shielded sources in vehicles loaded with scrap steel. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 315–318. Vienna: International Atomic Energy Agency, TECDOC-827.

Cox JW and Raymer LL (1976) The effect of potassium-salt muds on gamma-ray and spontaneous potential measurements. In: *Transactions of the 17th Annual Logging Symposium*, Society of Professional Well Log Analysts, Paper II, 19 pp.

Crew ME and Berkoff EW (1970) TWOPIT, a different approach to calibration of gamma ray logging equipment. *Log Analyst* 11: 26–32.

Czubek J (1971) Differential interpretation of gamma ray logs: I. Case of the static gamma ray curve. Report No. 760/1. Warsaw, Poland: Nuclear Energy Information Center, Polish Government Commissioner for Use of Nuclear Energy.

Czubek J (1972) Differential interpretation of gamma ray logs: II. Case of the dynamic gamma ray curve. Report No. 793/1. Warsaw, Poland: Nuclear Energy Information Center, Polish Government Commissioner for Use of Nuclear Energy.

Darnley AG (1970) Airborne gamma ray spectrometry. *Canadian Mining and Metallurgical Bulletin* 63: 145–154.

Darnley AG, Charbonneau BW, and Richardson KA (1977) The distribution of uranium in rocks as a guide to the recognition of uraniferous regions. In: *Recognition and Evaluation of Uraniferous Areas*, pp. 55–86. Vienna: IAEA.

Darnley AG and Ford KL (1989) Regional airborne gamma ray surveys: A review. In: Garland GD (ed.) *Proceedings of Exploration 87,* Special Volume 3, pp. 229–240. Toronto: Ontario Geological Survey.

Davis JD and Guilbert JM (1973) Distribution of radioelements potassium, uranium and thorium in selected porphyry copper deposits. *Economic Geology* 68: 145–160.

Delpierre ME (1982) Anomaly '60': A uraniferous granite pluton on Melville Peninsula, Northwest Territories. In: Maurice YT (ed.) *Uranium in Granites*, pp. 115–118. Ottawa, ON: Geological Survey of Canada, Paper 81-23.

Dickson BL and Scott KM (1997) Interpretation of aerial gamma ray surveys – Adding the geochemical factors. *AGSO Journal of Australian Geology and Geophysics* 17: 187–200.

Dodd PH and Eschliman DH (1972) Borehole logging techniques for uranium exploration and evaluation. In: Bowie SHU, Davis M, and Ostle D (eds.) *Uranium Prospecting Handbook*, pp. 244–276. London: Institution of Mining and Metallurgy.

Doig R (1968) The natural gamma ray flux: In situ analysis. *Geophysics* 33: 311–328.

Dredge LA, McMartin I, and Ford K (2005) Till geochemistry, gamma ray spectrometry and gold grain counts, Wager Bay, mainland Nunavut (NTS 56G). Open File 5014, Geological Survey of Canada, 184 p.

Dubé B and Gosselin P (2007) Greenstone-hosted quartz-carbonate vein deposits. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 49–73. St. John's, NL: Geological Association of Canada, Mineral Deposits, Division Special Publication No. 5.

Dubé B, Gosselin P, Mercier-Langevin P, Hannington M, and Galley A (2007) Gold-rich volcanogenic massive sulphide deposits. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 75–94. St. John's, NL: Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

Durocher ME (1983) The nature of hydrothermal alteration associated with the Madsen and Starratt-Olsen gold deposits, Red Lake area. In: Colvine AC (ed.) *The Geology of Gold in Ontario*, pp. 123–140. Sudbury, ON: Ontario Geological Survey, Miscellaneous Paper 110.

Duval JS Jr, Cook B, and Adams JAS (1971) Circle of investigation of an airborne gamma ray spectrometer. *Journal of Geophysical Research* 76: 8466–8470.

Dyck W (1969) Radon determination apparatus for geochemical prospecting for uranium. In: *Field and Laboratory Methods Used by the Geological Survey of Canada in Geochemical Surveys*, p. 30. Ottawa, ON: Geological Survey of Canada No. 10, Paper 68-21.

Evans RD (1955) *The Atomic Nucleus*. p. 972. New York: McGraw-Hill.

Fertl WH (1979) Gamma ray spectral data assists in complex formation evaluation. *Log Analyst* 20: 3–37.

Fertl WH and Rieke III HH (1979) Gamma ray spectral evaluation techniques identify fractured shale reservoir and source rock characteristics. *Society of Petroleum Engineers of AIME*, SPE 8454.

Fesq HW, Kable EJD, and Gurney JJ (1975) Aspects of the geochemistry of kimberlites from the Premier Mine and other South African occurrences, with particular reference to the rare earth elements. *Physics and Chemistry of the Earth* 9: 686–707.

Ford KL (1993) Radioelement mapping of parts of the Musquodoboit Batholith and Liscomb complex, Meguma zone Nova Scotia. In: Sangster AL (ed.) *Mineral Deposit Studies in Nova Scotia*, vol. 2, pp. 71–111. Ottawa, ON: Geological Survey of Canada, Paper 91-9.

Ford KL, Delabio RNW, and Rencz AN (1988) Geological, geophysical and geochemical studies around the Allan Lake carbonatite, Algonquin Park, Ontario. *Journal of Geochemical Exploration* 30: 99–121.

Ford KL, Harris JR, Shives R, Carson J, and Buckle J (2008a) Gamma ray spectrometry. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping*, pp. 53–97. Ottawa, ON: Geological Survey of Canada, ch. 4, Open File 5643.

Ford K, Harris JR, Shives R, Carson J, and Buckle J (2008b) Remote predictive mapping 2. Gamma ray spectrometry: A tool for mapping Canada's North. *Geoscience Canada* 35: 109–126.

Ford KL, Keating P, and Thomas MD (2007) Overview of geophysical signatures associated with Canadian ore deposits. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 939–970. St. John's, NL: Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

Ford KL and O'Reilly GA (1985) Airborne gamma ray spectrometric surveys as an indicator of granophile element specialization and associated mineral deposits in the granitic rocks of the Meguma Zone of Nova Scotia, Canada. In: *High Heat Production (HHP) Granites, Hydrothermal Circulation and Ore Genesis*, pp. 113–133. London, England: Institution of Mining and Metallurgy.

Ford KL, Savard M, Dessau JC, Pellerin E, Charbonneau BW, and Shives RBK (2001) The role of gamma ray spectrometry in radon risk evaluation: A case history from Oka, Quebec. *Geoscience Canada* 28: 59–64.

Gaboreau S, Vieillard P, Bruneton P, and Besufort D (2003) Aluminium phosphate sulphate minerals associated with unconformity uranium deposits of the Kombolgie Basin (Northern territory, Australia). In: *Proceedings of an International Conference on Uranium Geochemistry 2003*, Nancy-France, 13–16 April, Extended Abstracts, pp. 157–160.

Galley AG, Hannington MD, and Jonasson IR (2007) Volcanogenic massive sulphide deposits. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 141–161. St. John's, NL: Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

Gandhi SS, Prasad N, and Charbonneau BW (1996) Geological and geophysical signatures of a large polymetallic exploration target at Lou Lake, southern Great Bear magmatic zone, Northwest Territories. In: *Current Research 1996-E*, pp. 147–158. Ottawa, ON: Geological Survey of Canada.

Geological Survey of Canada (1980) Airborne gamma ray spectrometric map, Dihourse Lake, Quebec-Newfoundland, Geophysical Series Map 35124(08)G.

Geological Survey of Canada (1985) Airborne gamma ray spectrometric map, NTS 14D/05, Quebec-Newfoundland, Geophysical Series Map 35414(05)G.

Gillespie AR, Kahle AB, and Walker RE (1986) Color enhancement of highly correlated images. I. Decorrelation and HIS contrast stretches. *Remote Sensing of Environment* 20: 209–235.

Graham DF, Rencz AN, and Singhroy VH (1994) GSC – CCRS storefront project, conclusions from selected projects. Open File 2674, Geological Survey of Canada.

Grasty RL (1975) Uranium measurement by airborne gamma ray spectrometry. *Geophysics* 40: 503–519.

Grasty RL (1976a) The "field of view" of gamma ray detectors – A discussion. In: *Report of Activities, Part B*, pp. 81–82. Ottawa, ON: Geological Survey of Canada, Paper 76-1B.

Grasty RL (1976b) *A Calibration Procedure for an Airborne Gamma Ray Spectrometer*. Ottawa, ON: Geological Survey of Canada, Paper 76-16, pp. 1–9.

Grasty RL (1977) A general calibration procedure for airborne gamma ray spectrometers. In: *Report of Activities, Part C*, pp. 61–62. Ottawa, ON: Geological Survey of Canada, Paper 77-1C.

Grasty RL (1979) Gamma ray spectrometric methods in uranium exploration – Theory and operation procedures. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 147–162. Ottawa, ON: Geological Survey of Canada Economic Geology Report 31.

Grasty RL (1983) Radiation sources. *GEOS* 12: 1–5.

Grasty RL (1987) *The Design, Construction and Application of Airborne Gamma Ray Spectrometer Calibration Pads – Thailand*. p. 34. Ottawa, ON: Geological Survey of Canada, Paper 87-10.

Grasty RL (1995) Environmental monitoring by airborne gamma ray spectrometry, experience at the Geological Survey of Canada. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 93–102. Vienna: International Atomic Energy Agency, TECDOC-827.

Grasty RL and Blanchard Y (1993) *The Design, Construction and Application of Concrete Models for Calibrating Borehole Gamma Ray Spectrometers*. p. 14. Ottawa, ON: Geological Survey of Canada, Bulletin 454.

Grasty RL, Carson JM, Charbonneau BW, and Holman PB (1984) *Natural Background Radiation in Canada*. p. 39. Ottawa, ON: Geological Survey of Canada, Bulletin 360.

Grasty RL and Charbonneau BW (1974) Gamma ray spectrometer calibration facilities. In: *Report of Activities, Part B*, pp. 69–71. Ottawa, ON: Geological Survey of Canada Paper 74-IB.

Grasty RL and Cox JR (1997) A car-borne gamma ray spectrometer system for natural radioactivity mapping and environmental monitoring. *RESUME 95*, *Rapid Environmental Surveying Using Mobile Equipment*. Report, Nordic Nuclear Safety Research Secretariat, pp. 71–90.

Grasty RL and Darnley AG (1971) *The Calibration of Gamma Ray Spectrometers for Ground and Airborne Use*. p. 27. Ottawa, ON: Geological Survey of Canada, Paper 71–17.

Grasty RL, Holman PB, and Blanchard Y (1991) Transportable Calibration Pads for Ground and Airborne Gamma Ray Spectrometers, p. 25. Ottawa, ON: Geological Survey of Canada, Paper 90-23.

Grasty RL and Hovgaard J (1996) The calibration of upward detectors in gamma ray surveys. Expanded Abstracts, *Society of Exploration Geophysicists Annual Meeting*, Denver, pp. 1442–1445.

Grasty RL, Kosanke KL, and Foote RS (1979) Field of view of airborne gamma ray detectors. *Geophysics* 44: 1447–1457.

Grasty RL and Minty BRS (1995) *A Guide to the Technical Specifications for Airborne Gamma Ray Surveys*, p. 89 Canberra: Australian Geological Survey Organization, Record 1995/60.

Grasty RL and Sander L (2001) Airborne gamma ray surveys over Canadian nuclear sites. In: *63rd Annual Conference*, Expanded abstracts, European Association of Geoscientists and Engineers.

Grasty RL, Walters BRB, Hovgaard J, and LaMarre JR (2001) Calibration of a 7.6 cm $\times$ 7.6 cm (3 inch $\times$ 3 inch) sodium iodide gamma ray spectrometer for air kerma rate. *Radiation Protection Dosimetry* 94: 309–316.

Grasty RL, Wilkes PG, and Kooyman R (1988) *Background Measurements in Gamma Ray Surveys*. Ottawa, ON: Geological Survey of Canada, Paper 88-11.

Green AA, Berman M, Switzer P, and Craig MD (1988) A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing* GE-26: 65–74.

Gregory AF and Horwood JL (1961) A laboratory study of gamma ray spectra at the surface of rocks. Mines Branch Research Report R85. Ottawa, ON: Department of Energy, Mines and Resources.

Hansink JD (1976) Equilibrium analysis of a sandstone roll front uranium deposit. In: *Exploration for Uranium Ore Deposits*. *IAEA Proceedings Series*, pp. 683–693. Vienna: IAEA.

Harris JR (1989) Clustering of gamma ray spectrometer data using a computer image analysis system. In: Agterberg FP and Bonham-Carter GF (eds.) *Statistical Applications in the Earth Sciences*, pp. 19–31. Ottawa, ON: Geological Survey of Canada, Paper 89-9.

Harris JR, Bowie C, Rencz AN, and Graham D (1994) Computer enhancement techniques for the integration of remotely sensed, geophysical, and thematic data for the geosciences. *Canadian Journal of Remote Sensing* 20: 210–221.

Harris JR, Ford K, Charbonneau B, and Buckle J (2008) Application of gamma ray spectrometer data for lithological mapping in a Cordilleran environment, Sekwi Region, Northwest Territories. In: Harris JR (ed.) *Remote Predictive Mapping: An Aid for Northern Mapping, Case Study 5*, pp. 203–223. Ottawa, ON: Geological Survey of Canada.

Haycock M (1953) Memorial of Hardy Vincent Ellsworth. *American Mineralogist* 38: 427–431.

Henderson JR (1987) Geology, southeastern Melville Peninsula, District of Franklin, Northwest Territories. Geological Survey of Canada, Map 1655A, scale 1:100 000.

Hetu RJ, Holman PB, Charbonneau BW, Prasad N, and Gandhi SS (1994) Airborne geophysical survey, Mazenod Lake area, Northwest Territories, 1993 (NTS 85N/10 and parts of 85N/11, 14, 15). Open File 2806, Geological Survey of Canada.

Hetu R, Rangelov R, Grant JA, Grasty RL, and Graham D (1993) Monitoring natural and man-made radiation by airborne gamma ray spectrometry, Kozlodui Area, Bulgaria, 1990. Open File 2573, Geological Survey of Canada.

Holman PB, Carson JM, Ford KL, Grant JA, and Shives RBK (2001) Airborne gamma ray spectrometry compilation series, Quoich River, Nunavut. Open File 4005, scale 1:1,000,000, Geological Survey of Canada.

Horsfall KR (1997) Airborne magnetic and gamma ray data acquisition. *AGSO Journal of Australian Geology and Geophysics* 17: 159–174.

Hovgaard J and Grasty RL (1997) Reducing statistical noise in airborne gamma ray data through spectral component analysis. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth decennial Conference on Mineral Exploration*, pp. 753–764. Canada: GEO F/X.

Hyde EK, Perlman I, and Seaborg GT (1964) *The Nuclear Properties of the Heavy Elements: Detailed Radioactivity Properties*, vol. 2, p. 1107. New Jersey: Prentice Hall.

IAEA (1976) *Radiometric Reporting Methods and Calibration in Uranium Exploration*, p. 57. Vienna: International Atomic Energy Agency, Technical Reports Series No. 174.

IAEA (1977) Nuclear techniques and mineral resources. In: *Proceedings of the International Symposium on Nuclear Techniques in Exploration, Extraction and Processing of Mineral Resources*, p. 654. Vienna: International Atomic Energy Agency.

IAEA (1979) *Gamma Ray Surveys in Uranium Exploration*. p. 90. Vienna: International Atomic Energy Agency, Technical Reports Series No. 186.

IAEA (1982) Borehole Logging for Uranium Exploration – A Manual, p. 279. Vienna: International Atomic Energy Agency, Technical Report Series No. 212.

IAEA (1986) *Practical Borehole Logging Procedures for Mineral Exploration with Emphasis on Uranium*, p. 44. Vienna: International Atomic Energy Agency, Technical Report Series No. 259.

IAEA (1987) *Preparation and Certification of IAEA Gamma Ray Spectrometry Reference Materials RGU-1, RGTh-1 and RGK-1*, p. 48. Vienna: International Atomic Energy Agency Technical Report IAEA/RL/148.

IAEA (1989) *Construction and Use of Calibration Facilities for Radiometric Field Equipment*, p. 86 Vienna: International Atomic Energy Agency, Technical Reports Series No. 309.

IAEA (1991) *Airborne Gamma Ray Spectrometer Surveying*, p. 97 Vienna: International Atomic Energy Agency, Technical Reports Series No. 323.

IAEA (1995) *Application of Uranium Exploration Data and Techniques in Environmental Studies*, p. 351 Vienna: International Atomic Energy Agency, IAEA-TECDOC-827.

IAEA (2003) *Guidelines for Radioelement Mapping Using Gamma Ray Spectrometry Data*, p. 173. Vienna: International Atomic Energy Agency, IAEA TECDOC-1363.

IAEA (2010) *Radioelement Mapping*, p. 96 Vienna: International Atomic Energy Agency, IAEA Nuclear Energy Series No. NF-T-1.3.

IAEA (2013) *Advances in Airborne and Ground Geophysical Methods for Uranium Exploration*, p. 72. Vienna: International Atomic Energy Agency, IAEA Nuclear Energy Series No. NF-T-1.5.

Kaplan I (1964) *Nuclear Physics*. Reading, MA: Addison-Wesley Publishing Company.

Keppie JD (compiler) (2000) Geological map of the Province of Nova Scotia. Nova Scotia Department of Natural Resources, Minerals and Energy Branch, Map ME 2000-1, scale 1:500 000.

Khan MS (2012) *Source Localization Using a Directional Gamma Ray Spectrometer*. M. Sc. Thesis, Department of Physics, Carleton University, Ottawa, Canada, November 2012.

Killeen PG (1978) Gamma ray spectrometric calibration facilities: A preliminary report. In: *Current Research, Part A*, pp. 243–247. Ottawa, ON: Geological Survey of Canada, Paper 78-IA.

Killeen PG (1979) Gamma ray spectrometric methods in uranium exploration: Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, pp. 163–230. Ottawa, ON: Geological Survey of Canada, Economic Geology Report 31.

Killeen PG (1982) Gamma ray logging and interpretation. In: Fitch AA (ed.) *Developments in Geophysical Exploration Methods-3*, pp. 95–150. Essex: Applied Science Publishers, ch. 4.

Killeen PG (1986a) A system of deep test holes and calibration facilities for developing and testing new borehole geophysical techniques. In: Killeen PG (ed.) *Borehole Geophysics for Mining and Geotechnical Applications*, pp. 29–46. Ottawa, ON: Geological Survey of Canada, Paper 85-27.

Killeen PG (ed.) (1986b) *Borehole Geophysics for Mining and Geotechnical Applications*, p. 400. Ottawa, ON: Geological Survey of Canada, Paper 85-27.

Killeen PG (1997a) Borehole geophysics: Exploring the third dimension. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial Conference on Mineral Exploration*, pp. 31–42. Canada: GEO F/X.

Killeen PG (1997b) Nuclear techniques for ore grade estimation. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial Conference on Mineral Exploration*, pp. 677–684. Canada: GEO F/X.

Killeen PG (2013) Exploration trends and developments in 2012. *The Northern Miner* 99: 28, Supplement to 25 February 2013.

Killeen PG and Cameron GW (1977) Computation of in situ potassium, uranium and thorium concentration from portable gamma ray spectrometer data. In: *Report of Activities, Part A*, pp. 91–92. Ottawa, ON: Geological Survey of Canada, Paper 77-IA.

Killeen PG and Carmichael CM (1970) Gamma ray spectrometer calibration for field analysis of thorium, uranium and potassium. *Canadian Journal of Earth Science* 7: 1093–1098.

Killeen PG and Carmichael CM (1976) *Radioactive Disequilibrium Determinations, Part 1: Determination of Radioactive Disequilibrium in Uranium Ores by Alpha-Spectrometry*, pp. 1–18. Ottawa, ON: Geological Survey of Canada, Paper 75-38.
Killeen PG and Conaway JG (1978) New facilities for calibrating gamma ray spectrometric logging and surface exploration equipment. *Canadian Mining and Metallurgical Bulletin* 71: 84–87.
Killeen PG, Conaway JG, and Bristow Q (1978) A gamma ray spectral logging system including digital playback, with recommendations for a new generation system. In: *Current Research, Part A*, pp. 235–241. Ottawa, ON: Geological Survey of Canada, Paper 78-IA.
Killeen PG and Elliott BE (1990) Model boreholes for gamma ray logging: Intercalibration of North American and European calibration facilities. In: *Transactions of the 13th SPWLA European Formation Evaluation Symposium*, Budapest, 22–26 October 1990, paper GG, p. 14.
Killeen PG, Mwenifumbo CJ, and Elliott BE (1995) The Pseudo-geological log: Using geophysical logs as an aid to geological logging in volcanogenic massive sulphides. In: Research Current (ed.) *Part E*, pp. 321–330. Ottawa, ON: Geological Survey of Canada, Paper 95-1E.
Killeen PG, Mwenifumbo CJ, Elliott BE, and Chung CJ (1997) Improving exploration efficiency by predicting drill core logs with geophysical logs. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral Exploration*, pp. 713–716. Canada: GEO F/X.
Killeen PG and Richardson KA (1980) Regional radiogenic heat production mapping using airborne gamma ray spectrometry. In: *Current Research, Part B*, pp. 227–232. Ottawa, ON: Geological Survey of Canada, Paper 80-1B.
Knoll GF (1989) *Radiation Detection and Measurement*, 2nd edn. New York: Wiley.
Kresten P (1974) Uranium in kimberlites and associated rocks with special reference to Lesotho occurrences. *Lithos* 7: 171–180.
Lapp RE and Andrews HL (1963) *Nuclear Radiation Physics*. Englewood Cliffs, NJ: Prentice-Hall.
Lederer CM, Hollander JM, and Perlman I (1968) *Table of Isotopes*, 6th edn, p. 594. New York: Wiley.
Lederer CM and Shirley VS (eds.) (1978) *Table of Isotopes*, 7th edn. New York: Wiley.
Lee JB, Woodyatt AS, and Berman M (1990) Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform. *IEEE Transactions on Geoscience and Remote Sensing* 28: 295–304.
Lovborg L (1984) The calibration of portable and airborne gamma ray spectrometers – Theory, problems and facilities. Report Riso-M-2456. Riso National Laboratory, DK-4000 Roskilde, Denmark.
Lovborg L, Botter-Jensen L, and Kirkegaard P (1978) Experiences with concrete calibration sources for radiometric field instruments. *Geophysics* 43: 543–549.
Lovborg L, Grasty RL, and Kirkegaard P (1978) A guide to the calibration constants for aerial gamma ray surveys in geoexploration. In: *Aerial Techniques for Environmental Monitoring*, p. 193. New York: American Nuclear Society.
Lovborg L and Kirkegaard P (1974) Response of 3"×3" NaI(Tl) detectors to terrestrial gamma radiation. *Nuclear Instruments and Methods* 121: 239–251.
Lovborg L, Kirkegaard P, and Rose-Hansen J (1972) Quantitative interpretation of the gamma ray spectra from geologic formations. In: Adams JAS, Lowder WM, and Gesell TF (eds.) *The Natural Radiation Environment II*, pp. 155–180. Springfield, VA: US Department of Commerce.
Lovborg L, Kunzendorf H, and Hansen J (1969) Use of field gamma spectrometry in the exploration of uranium and thorium deposits in South Greenland. In: *Nuclear Techniques and Mineral Resources, IAEA Proceedings Series*, pp. 197–211. Vienna: IAEA.
Lovborg L, Wollenberg H, Sorensen P, and Hansen J (1971) Field determination of uranium and thorium, exemplified by measurements in the Ilimaussaq alkaline intrusion, South Greenland. *Economic Geology* 66: 368–384.
Lutts BG and Mineeva IG (1973) Uranium and thorium in Kimberlites of Siberia. *Geochemistry International (Geokhimiya)* 11: 1721–1724.
Macnae J (1979) Kimberlite and exploration geophysics. *Geophysics* 44: 1395–1416.
Madore C, Annesley I, and Wheatley K (2000) Petrogenesis, age, and uranium fertility of peraluminous leucogranites and pegmatites of the McClean Lake/Sue and Key Lake/P-Patch deposit areas, Saskatchewan. *Extended Abstracts in Geological Association of Canada-Mineralogical Association of Canada* (*GAC-MAC*) *Program 25*, 1041. PDF, p. 4.
Mahdavi A (1964) The thorium, uranium, and potassium contents of Atlantic and Gulf Coast beach sands. In: Adams JAS and Lowder WM (eds.) *The Natural Radiation Environment*, pp. 87–114. Chicago, IL: University of Chicago Press.
Mathews MA, Koizumi CJ, and Evans HB (1978) D.O.E. – Grand Junction logging model data synopsis. Report GJBX-76(78). US Department of Energy, Grand Junction Office.
Mechtly EA (1973) *The International System of Units, Physical Constants and Conversion Factors*. Washington, DC: National Aeronautics and Space Administration, NASA SP-7012, 2nd Revision.
Milligan PR and Gunn PJ (1997) Enhancement and presentation of airborne geophysical data. *AGSO Journal of Australian Geology and Geophysics* 17: 63–75.
Minty BRS (1992) Airborne gamma ray spectrometric background estimation using full spectrum analysis. *Geophysics* 57: 279–287.
Minty BRS (1997) Fundamentals of airborne gamma ray spectrometry. *AGSO Journal of Australian Geology and Geophysics* 17: 39–50.
Minty BRS (1998) Multichannel models for the estimation of radon background in airborne gamma ray spectrometry. *Geophysics* 63: 1986–1996.
Minty BRS (2000) Reducing noise in airborne gamma ray spectra. Australian Society of Exploration Geophysicists, Preview, Issue no. 89, 35–41.
Minty BRS (2003) Accurate noise reduction for airborne gamma ray spectrometry. *Exploration Geophysics* 34: 207–215.
Minty B, Luyendyk A, and Brodie R (1997) Calibration and data processing for airborne gamma ray spectrometry. *AGSO Journal of Australian Geology and Geophysics* 17: 51–62.
Minty BRS and McFadden P (1998) Improved NASVD smoothing of airborne gamma ray spectra. *Exploration Geophysics* 29: 516–523.
Morse RH (1976) Radon counters in uranium exploration. In: *Exploration for Uranium Ore Deposits*. *IAEA Proceedings Series*, pp. 229–239. Vienna: IAEA.
Morse JG (ed.) (1977) *Developments in Economic Geology, 7: Nuclear Methods in Mineral Exploration and Production*, pp. 280. Amsterdam: Elsevier Scientific Publishing Company.
Mwenifumbo CJ (1993) Kernel density estimation in the analysis and presentation of borehole geophysical data. *Log Analyst* 34: 34–45.
Mwenifumbo CJ and Bernius GR (2007) Crandallite-group minerals: Host of thorium enrichment in the eastern Athabasca Basin, Saskatchewan. In: Jefferson CW and Delaney G (eds.) *EXTECH IV: Geology and Uranium EXploration TECHnology of the Proterozoic Athabasca Basin*, pp. 521–532. Ottawa, ON: Geological Survey of Canada, Bulletin 588.
Mwenifumbo CJ, Elliott BE, Salisbury M, Jefferson CW, Bernius GR, and Pflug KA (2004) Physical rock properties from the Athabasca Group: Designing geophysical exploration models for unconformity uranium deposits. *Journal of Applied Geophysics* 55: 117–135.
Mwenifumbo CJ, Hunter JAM, and Killeen PG (1996) Geophysical characteristics of Canadian kimberlites. In: LeCheminant AN, Richardson DG, DiLabio RNW, and Richardson KA (eds.) *Searching for Diamond in Canada*, pp. 237–240. Ottawa, ON: Geological Survey of Canada, Open File 3228.
Mwenifumbo CJ, Kerswill J, Elliott BE, Falck H, Thompson P, and Ootes L (2006) Geophysical characteristics of gold mineralization at the Giant and Crestaurum deposits, Yellowknife greenstone belt, Northwest Territories. In: Anglin CD, Falck H, Wright DF, and Ambrose EJ (eds.) *Gold in the Yellowknife Greenstone Belt, Northwest Territories: Results of the EXTECH III Multidisciplinary Research Project*, pp. 340–354. St. John's, NL: Geological Association of Canada, Special Paper No. 3.
Mwenifumbo CJ and Killeen PG (1986) Natural gamma ray logging in volcanic rocks: The Mudhole and Clementine base metal prospects. In: Kirkham RV (ed.) *Buchans Geology, Newfoundland: 1982–1984*, pp. 263–272. Ottawa, ON: Geological Survey of Canada, Paper 86-24.
Mwenifumbo CJ and Kjarsgaard BA (2000) Gamma ray logging and radioelement distribution in the Fort à la Corne pipe 169. *Exploration and Mining Geology, Journal of the Geological Society of the Canadian Institute of Mining* 8: 133–147.
Mwenifumbo CJ, Urbancic TI, and Killeen PG (1983) Preliminary studies in gamma ray spectral logging in exploration for gold. In: *Current Research, Part A*, pp. 391–397. Ottawa, ON: Geological Survey of Canada, Paper 83-1A.
Nargolwalla SS (1975) In-situ borehole logging of geologic materials by neutron-capture gamma ray measurement. *Transactions of the American Nuclear Society* 21: 107.
Nordic Nuclear Safety Research (NKS) (1997) RESUME 95, Rapid environmental surveying using mobile equipment, report. Nordic Nuclear Safety Research Secretariat, p. 440. ISBN 87-7893-014-6.
NURE (1976) National uranium resource evaluation annual NURE report. Grand Junction Office Report GJBX-II (77). US Energy Research and Development Administration.
Ontario Geological Survey (1991) Bedrock geology of Ontario, southern sheet. Ontario Geological Survey, Map 2544, scale 1:1,000,000.
Osterlund SE (1986) A new test site for the calibration of geophysical equipment. In: Killeen PG (ed.) *Borehole Geophysics for Mining and Geotechnical Applications*, pp. 53–55. Ottawa, ON: Geological Survey of Canada, Paper 85-27.

Ostrihansky L (1976) *Radioactive Disequilibrium Determinations, Part 2: Radioactive Disequilibrium Investigations, Elliot Lake area, Ontario*. Ottawa, ON: Geological Survey of Canada, Paper 75-38, pp. 19–48.

Otton JK, Gundersen LCS, Schumann RR, Reimer GM, and Duval JS (1995) Uranium resource assessment and exploration data for geologic radon potential assessment in the United States. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 135–138. Vienna: International Atomic Energy Agency, TECDOC-827.

Paul DK, Gale NH, and Harris PG (1977) Uranium and thorium abundances in Indian kimberlites. *Geochimica et Cosmochimica Acta* 41: 335–339.

Pflug KA, Killeen PG, and Mwenifumbo CJ (1993) Application of spectral gamma ray logging to low-level radioactive waste management studies. In: *Proceedings of the 5th International Symposium of the Minerals and Geotechnical Logging Society*, Tulsa, 24–28 October 1993, Paper V, pp. 1–18.

Pflug KA, Killeen PG, and Mwenifumbo CJ (1996) Borehole geophysical logs in gold deposits in the Kirkland Lake area, Ontario (Macassa Gold Mine and Victoria Creek Gold Deposit). Open File 3153, Geological Survey of Canada.

Portnov AM (1987) Specialization of rocks towards potassium and thorium in relation to mineralization. *International Geology Review* 29: 326–344.

Rhodes DF and Mott WE (1966) Quantitative interpretation of gamma ray spectral logs. *Geophysics* 31: 410–418.

Richardson KA (1964) Thorium, uranium, and potassium in the Conway granite, New Hampshire, USA. In: Adams JAS and Lowder WM (eds.) *The Natural Radiation Environment*, pp. 39–50. Chicago, IL: University of Chicago Press.

Rosholt JN Jr (1959) *Natural Radioactive Disequilibrium of the Uranium Series. U.S. Geological Survey, Bulletin 1084-A*, pp. 30. Washington: U.S. Govt. Printing Office.

Sanderson DCW, Allyson JD, and Tyler AN (1995) Rapid quantification and mapping of radiometric data for anthropogenic and technologically enhanced natural nuclides. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 197–216. Vienna: International Atomic Energy Agency, TECDOC-827.

Sanderson DCW, Allyson JD, Tyler AN, and Scott EM (1995) Environmental applications of airborne gamma spectrometry. In: *Application of Uranium Exploration Data and Techniques in Environmental Studies*, pp. 71–92. Vienna: International Atomic Energy Agency, TECDOC-827.

Sanderson DCW and Ferguson JM (1997) The European capability for environmental airborne gamma ray spectrometry. *Radiation Protection Dosimetry* 73: 213–218.

Sass J, Killeen PG, and Mustonen E (1968) Heat flow and surface radioactivity in the Quirke Lake Syncline near Elliot Lake, Ontario. *Canadian Journal of Earth Sciences* 5: 1417–1428.

Schetselaar EM (2000) *Integrated Analyses of Granite – Gneiss Terrain from Field and Multisource Remotely Sensed Data: A Case Study from the Canadian Shield*. Enschede: ITC90-6164-180-2, ITC Dissertation 73, 269 p.

Schetselaar EM, Chung CF, and Kim K (2000) Classification of bedrock units in vegetated granite – Gneiss terrain by integration of airborne geophysical images and primary field data. *Remote Sensing of Environment* 71: 89–105.

Schetselaar EM, Harris JR, Lynds T, and de Kemp EA (2007) Remote predictive mapping 1. Remote predictive mapping (RPM): A strategy for geological mapping of Canada's North. *Geoscience Canada* 34: 93–111.

Schetselaar EM, Tiainen M, and Woldai T (2008) *Integrated Geological Interpretation of Remotely Sensed Data to Support Geological Mapping in Mozambique*, pp. 1–31. Espoo: Geological Survey of Finland, Special Paper 48.

Scott JH (1962) The GAMLOG computer program. US Atomic Energy Commission Report RME-143, Grand Junction, Colorado.

Scott JH (1963) Computer analysis of gamma ray logs. *Geophysics* 28: 457–465.

Scott JH and Dodd PH (1960) Gamma-only assaying for disequilibrium corrections. US Atomic Energy Commission Report RME-135, pp. 1–20.

Scott JH, Dodd PH, Droullard RF, and Mudra PJ (1961) Quantitative interpretation of gamma ray logs. *Geophysics* 26: 182–191.

Senftle FE and Hoyte AF (1966) Mineral exploration and soil analysis using in situ neutron activation. *Nuclear Instruments and Methods* 42: 93–103.

Serra O, Baldwin J, and Quirein J (1980) Theory, interpretation and practical applications of natural gamma ray spectroscopy. In: *Transactions of Society of Petroleum Well Log Analysts* (*SPWLA*) *21st Annual Logging Symposium*, Paper Q, pp. 1–30.

Shilts WW (1976) Glacial till and mineral exploration. In: Leggett RF (ed.) *Glacial Till: An Inter-disciplinary Study*, pp. 205-224. Royal Society of Canada, Special Publication 12, pp. 205–224.

Shives RBK, Charbonneau BW, and Ford KL (1997) The detection of potassic alteration by gamma ray spectrometry – Recognition of alteration related to mineralization. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial Conference on Mineral Exploration*, pp. 741–752. Canada: GEO F/X.

Shives RBK, Ford KL, and Charbonneau BW (1995) Geological Survey of Canada Workshop Manual: Applications of gamma ray spectrometric/magnetic/VLF-EM surveys, 85 p. Open File 3061, Geological Survey of Canada.

Siegbahn K (1968) *Alpha-, Beta-, and Gamma Ray Spectroscopy*, p. 1742. Amsterdam: North Holland Publishing Company.

Sinclair WD (2007) Porphyry deposits. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 223–243. St. John's, NL: Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5.

Smethurst MA, Strand T, Sundal AV, and Rudjord AL (2008) Large-scale radon hazard evaluation in the Oslofjord region of Norway utilizing indoor radon concentrations, airborne gamma ray spectrometry and geological mapping. *Science of the Total Environment* 407: 379–393.

Smith RJ (2002) Geophysics of iron-oxide copper-gold deposits. In: Porter TM (ed.) *Hydrothermal Iron Oxide Copper-Gold and Related Deposits: A Global Perspective*, vol. 2, pp. 357–367. Adelaide: PGC Publishing.

Smith AR and Wollenberg HA (1972) High-resolution gamma ray spectrometry for laboratory analysis of the uranium and thorium decay series. In: Adams JAS, Lowder WM, and Gesell TF (eds.) *The Natural Radiation Environment II*, pp. 181–231. Springfield, VA: US Department of Commerce.

Tanner AB (1964) Radon migration in the ground: A review. In: Adams JAS and Lowder WM (eds.) *The Natural Radiation Environment*, pp. 161–190. Chicago, IL: University of Chicago Press.

Tanner AB (1980) Radon migration in the ground: A supplementary review. In: Gesell TF and Lowder WM (eds.) *Natural Radiation Environment III*, vol. 1, pp. 5–56. Washington, DC: US Department of Energy Report CONF-780422.

Tanner AB, Moxham RM, and Senftle FE (1977) Assay for uranium and determination of disequilibrium by means of in situ high resolution gamma ray spectrometry. Open File Report 77-571. US Geological Survey.

Thomas MD, Ford KL, and Keating P (2011) Exploration geophysics for intrusion-hosted rare earth metals. Open File 6828 (poster). Geological Survey of Canada. doi:10.4095/288092.

Toens PD, van As D, and Vleggaar CM (1973) A facility at the national nuclear research center, Pelindaba, for the calibration of gamma-survey meters used in uranium prospecting operations. *Journal of the South Africa Institute of Mining and Metallurgy* 73: 428.

Tsoulfanidis N and Landsberger S (2010) *Measurement and Detection of Radiation*, 3rd edn. New York: CRC Press.

Urbancic TL and Mwenifumbo CJ (1986) Multiparameter logging techniques applied to gold exploration. In: Killeen PG (ed.) *Borehole Geophysics for Mining and Geotechnical Applications*, pp. 13–28. Ottawa, ON: Geological Survey of Canada, Paper 85-27.

Ward DL (1978) *Construction of the Calibration Pads Facility, Walker Field, Grand Junction, Colorado*. Washington, DC: US Department of Energy, Grand Junction Office Report GJBX-37(78).

Webster SS (1984) Comments on the use of gamma ray spectrometry for tin prospecting. *Exploration Geophysics* 15: 61–63.

West FG and Laughlin AW (1976) Spectral gamma logging in crystalline basement rocks. *Geology* 4: 617–618.

White SW, Gundersen, LCS, and Schumann, RR (1993) Development of EPA's map of radon zones. In: *Proceedings of the 1992 International Symposium on Radon and Radon Reduction Technology*, vol. 2. Symposium Oral Papers, Technical Sessions VII–XII: Research Triangle Park, North Carolina, US Environmental Protection Agency Report EPA-600/R-93/083b, pp. 8-51–8-86.

Wilford JR, Bierwirth PN, and Craig MA (1997) Application of airborne gamma ray spectrometry in soil/regolith mapping and applied geomorphology. *AGSO Journal of Australian Geology and Geophysics* 17: 201–216.

Wilson JA (1985) Crandallite group minerals in the Helikian Athabasca Group in Alberta, Canada. *Canadian Journal of Earth Sciences* 22: 637–641.

Wollenberg HA and Smith AR (1964) Studies in terrestrial gamma radiation. In: Adams JAS and Lowder WM (eds.) *The Natural Radiation Environment*, pp. 513–566. Chicago, IL: University of Chicago Press.

World Health Organization (2009) Fact Sheet No. 291: Radon and cancer, September 2009. Available at http://www.who.int/mediacentre/factsheets/fs291/en.

Wormold MR and Clayton CG (1976) Observations on the accuracy of gamma spectrometry in uranium prospecting. In: *Exploration for Uranium Ore Deposits*, IAEA Proceedings Series, pp. 149–172. Vienna: IAEA.

Yeates AN, Wyatt BW, and Tucker DH (1982) Application of gamma ray spectrometry to prospecting for tin and tungsten granites, particularly within the Lachlan Fold Belt, New South Wales. *Economic Geology* 77: 1725–1738.

11.15 Mineral Resources

ME Best, Bemex Consulting International, Victoria, BC, Canada

© 2015 Elsevier B.V. All rights reserved.

11.15.1	**Introduction**	525
11.15.2	**Metals and Their Minerals of Economic Interest**	526
11.15.3	**Physical Properties of Minerals and Rocks**	526
11.15.3.1	Magnetic Susceptibility of Rocks and Minerals	526
11.15.3.2	Mass Density of Rocks and Minerals	527
11.15.3.3	Resistivity of Rocks and Minerals	528
11.15.3.4	Seismic Velocity of Rocks and Minerals	528
11.15.3.5	Gamma Radiation from Rock and Minerals	529
11.15.3.6	Dielectric Constant of Rocks and Sediments	529
11.15.4	**Geophysical Methods Applied to Mineral Exploration**	530
11.15.4.1	Geophysical Applications to Regional/Geologic Mapping	531
11.15.4.2	Geophysical Applications by (Metal) Deposit Type	532
11.15.4.2.1	VMS deposits	533
11.15.4.2.2	Porphyry deposits	535
11.15.4.2.3	Kimberlite (diamonds) deposits	537
11.15.4.2.4	Unconformity-related uranium deposits	539
11.15.4.2.5	Laterite (nickel) deposits	544
11.15.4.2.6	Placer deposits	545
11.15.4.2.7	Other deposits	546
11.15.5	**Nonmetallic Deposits**	548
11.15.6	**Recent Geophysical Developments**	551
11.15.6.1	Airborne Gravity and Gravity Gradiometry	551
11.15.6.2	Land-Based Seismic Methods in Mineral Exploration	552
11.15.7	**Future Developments**	553
References		553

11.15.1 Introduction

Minerals provide many of the essential commodities necessary for society. Important metals such as iron, aluminum, copper, lead, zinc, molybdenum, nickel, and uranium are derived from metallic minerals. Nonmetallic minerals are equally important. Potash and phosphate for fertilizers, coal for thermal electricity and steel production, and cement for construction are common examples. Most minerals are found in low concentrations everywhere within the Earth's upper crust. However, the extraction of a mineral depends on that mineral being available in sufficient concentration to be mined economically. Such economic concentrations of minerals are referred to as mineral deposits. Mineral deposits are not randomly distributed within the Earth but generally occur in specific geologic environments. Successful exploration for a mineral deposit therefore requires knowledge of the relationship between the deposit and its geologic environment.

Modern mineral exploration requires a complex, multidisciplinary approach to be successful. Geology, geophysics, geochemistry, and drilling are examples of the disciplines required to effectively explore and develop a mineral deposit. Understanding mineral deposit models and their relation to geologic environments is necessary in order to focus exploration activity. There is no sense spending exploration funds on expensive geophysical and geochemical surveys for a mineral deposit that is in an area that does not have the appropriate geologic setting. Although geophysical methods are generally applied to detailed mapping, they also provide useful regional information to help geologists determine if an area has mineral potential or not.

The focus of this chapter is on the application of geophysical methods to mineral exploration and exploitation, not on theory and methodology, which can be found in other chapters of this volume. Several review articles on the state of the art in mining geophysics published during the last two decades (Lowe et al., 1999; Nabighian and Asten, 2002; Vallee et al., 2011) provide additional information. This chapter provides the reader with an up-to-date synopsis of the role of geophysics in the search for and extraction of economic minerals.

This chapter begins with a short discussion on the relationship between metals, minerals, and rocks and then proceeds with a discussion of the physical properties of rocks and minerals. A section follows on the application of geophysical methods to regional geologic mapping and a discussion of geophysical methods that are used for the direct detection of mineral deposits. The next section provides examples of the different types of geophysical responses that are associated with specific metal deposit types. This is followed by a section on the application of geophysical methods to nonmetallic mineral deposits. An up-to-date review of two significant recent geophysical developments that are now being applied

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00200-1

to mineral exploration and development is presented. This chapter ends with a section speculating on future developments and applications of geophysical methods to mineral exploration.

11.15.2 Metals and Their Minerals of Economic Interest

Most metals are not found as pure substances in nature but are combined with other elements to form minerals. Not all minerals containing metals are of economic interest because of the difficulty in extracting the metal from them. Metallic minerals are minerals that can be processed (usually by melting) to obtain metal products. Rocks are formed from combinations of different minerals.

Table 1 lists a number of important metals and their main economic minerals. The economic minerals are grouped according to whether they are oxides, sulfides, precious metals, or nonmetallic. Iron, aluminum, tin, tungsten, and uranium minerals are oxides, while the minerals of copper, lead, zinc, nickel, molybdenum, and silver are sulfide minerals. Lateritic nickel is an exception since the minerals associated with it are not sulfides but a mixture of hydrous and other minerals. Silver is a precious metal and could be placed in that group as well. Platinum group elements (PGEs) are often associated with sulfides, so they could equally well be grouped with sulfide minerals. Gold and diamonds are unique since they are mined directly as pure substances. Gold is also a by-product associated with several of the sulfide minerals listed in Table 1.

There are often minor mineral constituents associated with a given mineral deposit. For example, the main mineral associated with a volcanogenic massive sulfide (VMS) copper deposit is chalcopyrite, but there can be minor amounts of bornite, chalcocite (Cu_2S), and covallite (CuS). In addition, there can be nonsulfide copper minerals such as malachite ($Cu_3Co_3(OH)_2$) and azurite ($Cu_3(Co_3)_2\ (OH)_2$). There is usually a significant amount of iron sulfide, that is, pyrite (FeS_2). Another iron sulfide mineral that is often associated with VMS deposits is pyrrhotite ($Fe_{1-x}S_{x=0-0.2}$).

Nonmetallic minerals are treated separately in Section 11.15.6 because the geophysical methods used to explore for them are often different from those used to explore for metals. Diamonds, although they are nonmetal, are included in Table 1 because the geophysical techniques used to explore for them are similar to those used for the other minerals listed in the table.

Table 1 Important metals and their economic minerals

Metal	*Economic minerals*	*Comments*
	Oxide minerals	
Iron (Fe)	Magnetite (Fe_3O_4) Hematite (Fe_2O_3)	Magnetite more common
Tin (Sn)	Cassiterite (SnO_2)	
Uranium (U)	Pitchblende (UO_3, U_2O_5) Uraninite (UO_2)	Pitchblende more common
Tungsten (W)	Wolframite (Fe,Mn) WO_4 Scheelite ($CaWO_4$)	
Aluminum (Al)	Bauxite[a]	Main source of aluminum – consists of many minerals[a]
	Sulfide minerals	
Copper (Cu)	Chalcopyrite ($CuFeS_2$) Bornite (Cu_5FeS_4)	Bornite is more common in porphyry deposits
Lead (Pb)	Galena (PbS)	
Zinc (Zn)	Sphalerite (Zn,Fe)S	
Silver (Ag)	Argentite (Ag_2S)	Main ore of silver – also considered a precious metal
Nickel (Ni)	Pentlandite $(Fe,Ni)_9S_8$ Garnierite (hydrous Ni–Mg silicate) Goethite (FeO(OH))	Nickel laterites (garnierite, goethite) are not sulfides. There are other minerals for lateritic nickel as well
Molybdenum (Mo)	Molybdenite (MoS_2)	
	Precious metals	
Gold (Au)	Arsenopyrite (FeAsS) Native	Sulfide Pure metal
Platinum group elements (PGE)	(Pt, Pd)S Cooperite (Pt,Pd,Ni)S	Pt occurs with other minerals as well (Pt considered a precious metal like gold)
	Nonmetallic minerals	
Diamonds	Native	Not a metal but of economic interest (found in kimberlites)

[a]Bauxite is an aluminum ore and is the main source of aluminum. It consists mostly of the hydrous minerals gibbsite $Al(OH)_3$, boehmite γ-AlO(OH), and diaspore α-AlO(OH), in a mixture with the two iron oxides goethite and hematite, the clay mineral kaolinite, and small amounts of anatase TiO_2.

11.15.3 Physical Properties of Minerals and Rocks

There must be a contrast in the physical properties between a mineral deposit and the surrounding rock to be detectable using geophysical methods. The physical properties are (mass) density, magnetic susceptibility, velocity, resistivity (conductivity = 1/resistivity), chargeability, dielectric constant (electric permittivity), and gamma radiation. These are listed in Table 2 along with the corresponding geophysical method used to measure them. A short discussion of each of these physical properties follows.

11.15.3.1 Magnetic Susceptibility of Rocks and Minerals

The magnetic field measured by a magnetometer is the vector sum of the magnetic field caused by the induced magnetism in a body generated by the Earth's magnetic field and the remanent magnetism inherent within a magnetic body. The relationship between the induced magnetism in a body ($\mathbf{M}$) and the inducing field ($\mathbf{H}$), the Earth's magnetic field, is

$$\mathbf{M} = k\mathbf{H}$$

where k is called the magnetic susceptibility. The induced magnetism $\mathbf{M}$ and the inducing field $\mathbf{H}$ have units of amperes per meter (A m^{-1}), so k is unitless. The electromagnetic system of units (emu) for k and the SI system of units for k (note that

Table 2 Physical properties of rocks and minerals and the corresponding geophysical technique

Physical property	*Symbol*	*Units*	*Geophysical method*
Velocity	v	$m\ s^{-1}$	Reflection and refraction seismic
Acoustic impedance	ρv	$kg\ m^{-2}s^{-1}$	Reflection seismic
Density	ρ	$kg\ m^{-3}$	Gravity
Magnetic susceptibility	k	SI[a]	Magnetic
Resistivity	ρ	Ω-m	DC resistivity, electrical resistivity tomography (ERT)
Chargeability	m	$ms\ or\ mV\ V^{-1}$	Induced polarization (IP)
Conductivity[b]	σ	$S\ m^{-1}$	Electromagnetic
Gamma radiation		$Counts\ s^{-1}$	Radiometrics
Dielectric constant	K	No unit	Ground-penetrating radar (GPR)

[a]The SI unit is defined within the magnetic properties section in the succeeding text and is given as $k(\mathrm{SI}) = 4\pi \times k(\mathrm{emu})$ – electromagnetic unit that is a variation of the cgs system of units) where k is the magnetic susceptibility, that is, $\mathbf{M} = k\mathbf{H}$ where $\mathbf{M}$ is the induced magnetism and $\mathbf{H}$ is the Earth's magnetic field (both in units of $A\ m^{-1}$ so k is unitless).
[b]Resistivity (ρ) and conductivity (σ) are the reciprocal of each other ($\rho = 1/\sigma$). S(siemens) $= 1\ \Omega^{-1}$.

the centimeter–gram–second (cgs) system of units and the emu are equivalent) are related by the following expression:

$$k(\mathrm{SI}) = 4\pi \times k(\mathrm{emu})$$

Consequently, k(emu) is also unitless. From the aforementioned discussion, the magnetic susceptibility of minerals determines their ability to become magnetic under the influence of a magnetizing field. A related property is magnetic permeability (μ), which is related to the volumetric ability of materials to allow magnetic lines of force to pass through them. The relation between k and μ is given in the following equation:

$$\mu = \mu_o(1 + k)$$

where μ_o is the magnetic permeability of free space equal to $4\pi \times 10^{-7}\ H\,m^{-1}$. **Table 3** from Telford et al. (1976) provides typical ranges and average values of susceptibility (in emu units) for a variety of rock types. Note that there is a large range of values for these rocks.

There are only a few magnetic minerals that possess high magnetic susceptibility. Magnetite is the most important one, in terms of both its effect and distribution. Two other magnetic minerals of lesser frequency of occurrence and importance are pyrrhotite and ilmenite. Although man-made materials have much larger magnetic susceptibility values, these three minerals determine most of the magnetic properties of rocks. The volume percent of magnetite, both constituent and free, in rocks most often determines their magnetic susceptibility. **Figure 1** shows that percent magnetite and magnetic susceptibility of a rock follow a close linear relationship. Igneous rocks, especially mafic varieties, can contain magnetite in significant enough concentrations to cause localized and regional distortions in the Earth's magnetic field. Sedimentary rocks normally have very little or no magnetic mineral content; hence, their magnetic susceptibilities are very low and are normally magnetically quiet. Rock susceptibility is a highly variable parameter, even within a given rock unit (**Table 3**). The bulk susceptibility of a rock formation or body depends on the distribution of the individual grains of magnetite, ilmenite, and pyrrhotite in the rock matrix; this distribution can be highly variable.

Further details on magnetic properties and methods can be found in **Chapters 11.05** and **11.11**.

Table 3 Magnetic susceptibility of various rocks and minerals (susceptibility[a] $\times 10^{-6}$ (emu))

Type	*From Telford et al. (1976)*	*From Dobrin (1960)*	*From Parasinis (1975)*
Sedimentary			
Dolomite	0–75	14	–
Limestone	2–280	–	12.5–24 700
Sandstone	0–1660	16.9	–
Shale	5–1480	–	–
Metamorphic			
Amphibolite	–	–	–
Schist	25–260	–	–
Phyllite	–	–	–
Gneiss	10–2000	–	–
Quartzite	–	–	–
Serpentine	250–1400	14 000	–
Slate	0–3000	–	–
Igneous			
Granite	0–4000	28–2700	0–49 500
Rhyolite	20–3000	–	250–9500
Dolerite	100–3000	–	–
Augite syenite	2700–3000	–	–
Olivine–diabase	–	2000	–
Diabase	80–13 000	78–1050	10 000–150 000
Porphyry	20–16 700	47	225–210 000
Gabbro	80–7200	68.1–2370	–
Basalt	20–14 500	680	250–105 000
Diorite	50–10 000	46.8	–
Pyroxenite	–	–	–
Peridotite	7600–15 600	12 500	–
Andesite	–	–	
Minerals			
Magnetite		300 000–800 000	–
Pyrrhotite		125 000	5300
Ilmenite		135 000	–

[a]Ranges where available.

11.15.3.2 Mass Density of Rocks and Minerals

Approximate density ranges of rocks and unconsolidated sediments are given in **Table 4**. Further information on density of

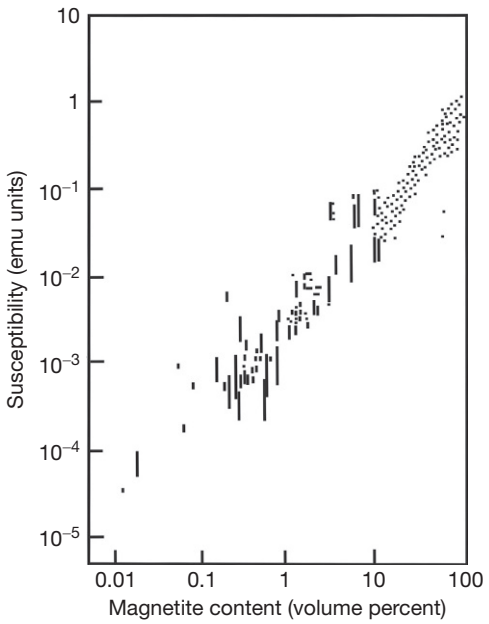


Figure 1 Relationship between magnetite susceptibility and magnetite content for a variety of rocks and ores. Modified from Lindsay DH, Andreasen GE, and Balsley JR (1966) Magnetic properties of rocks and minerals. *Handbook of Physical Constants* (revised edn.), Section 25, Geological Society of America, Memoir 95, p. 543. Page 46 of 46.

igneous and metamorphic rocks can be found in Table 1 of Salisbury et al. (2003). The bulk density of a rock depends on the different types of minerals within the rock matrix, porosity, and the fluid (water or air) filling the voids. An expression for the bulk density (ρ_b) of a rock or sediment is

$$\rho_b = \rho_m(1-\phi) + \rho_f\phi$$

where ρ_m is the matrix density, ρ_f is the fluid density, and ϕ is the porosity.

Clastic sedimentary rocks (sandstone, shale, and conglomerate), evaporites (halite, gypsum, and chalk), and coal generally have lower density than carbonate rocks (limestone and dolomite). The density of igneous rocks with minor amounts of iron-rich minerals (e.g., granite, granodiorite, gneiss, rhyolite, and quartzite, which are rich in quartz and feldspar) is generally lower than igneous rocks with significant amounts of iron-rich minerals (e.g., gabbro, basalt, amphibolite, diabase, and diorite). The range of density values for all the materials in Table 4 is narrow compared with the range of values for magnetic susceptibility and resistivity. Massive sulfide ore bodies generally have a much higher density than the host rock surrounding it. See Chapter 11.12 for additional information on density and the gravitational method (this volume).

11.15.3.3 Resistivity of Rocks and Minerals

Figure 2 shows typical resistivity ranges (Ω-m) and conductivity (which is the reciprocal of resistivity) ranges (mS m^{-1}) of rocks and sediments as well as massive sulfide mineralization, graphite, and water. There is a large range of resistivity values compared with other physical parameters. Electromagnetic

Table 4 Approximate density ranges of rocks and unconsolidated sediments

Type	*Density (kg m^{-3})*
Unconsolidated sediments[a]	
Alluvium (wet)	1960–2000
Clay	1630–2600
Sedimentary rocks[a]	
Shale	2060–2660
Sandstone	
Cretaceous	2050–2350
Triassic	2250–2300
Carboniferous	2350–2550
Limestone	2600–2800
Chalk	1940–2230
Dolomite	2280–2900
Halite	2100–2400
Igneous rocks[a]	
Granite	2520–2750
Granodiorite	2670–2790
Anorthosite	2610–2750
Basalt	2700–3200
Gabbro	2850–3120
Gneiss	2610–2990
Quartzite	2600–2700
Amphibolite	2790–3140
Massive sulfides[b]	
Pyrite	4910
Sphalerite	4070
Chalcopyrite	4280
Pyrrhotite	4710
Pentlandite	4680
Bornite	5070
Galena	7510

[a]Modified from Kearey and Brooks (1984).
[b]Modified from Salisbury et al. (2003).

(EM) responses from a conductive target are significantly stronger than the EM responses from a resistive target. This is one of the reasons why EM methods are so effective for massive sulfide exploration. Unweathered igneous and metamorphic crystalline rocks are the most resistive rocks, followed by igneous and metamorphic weathered rocks. Carbonate and conglomerate rocks are the most resistive of the sedimentary rocks and shale the most conductive. Unconsolidated sediments, such as sand and gravel, can have resistivity values that can vary over many decades depending on porosity, water conductivity, and clay content. See Sumner (1976, pp. 32–37) and Keller and Frischknecht (1966) for more details on the resistivity of rocks and minerals. Additional details on resistivity and electrical methods can be found in Chapter 11.04 and Chapter 11.08.

11.15.3.4 Seismic Velocity of Rocks and Minerals

Approximate ranges for the *P*-wave (pressure-wave) velocity of rocks and unconsolidated sediments are given in Table 5. Further information on *P*-wave velocity and *S*-wave (shear-wave) velocity of igneous and metamorphic rocks can be found in Table 1 of Salisbury et al. (2003). Unconsolidated sediments generally have *P*-wave velocities <2000 m s^{-1}. Sedimentary

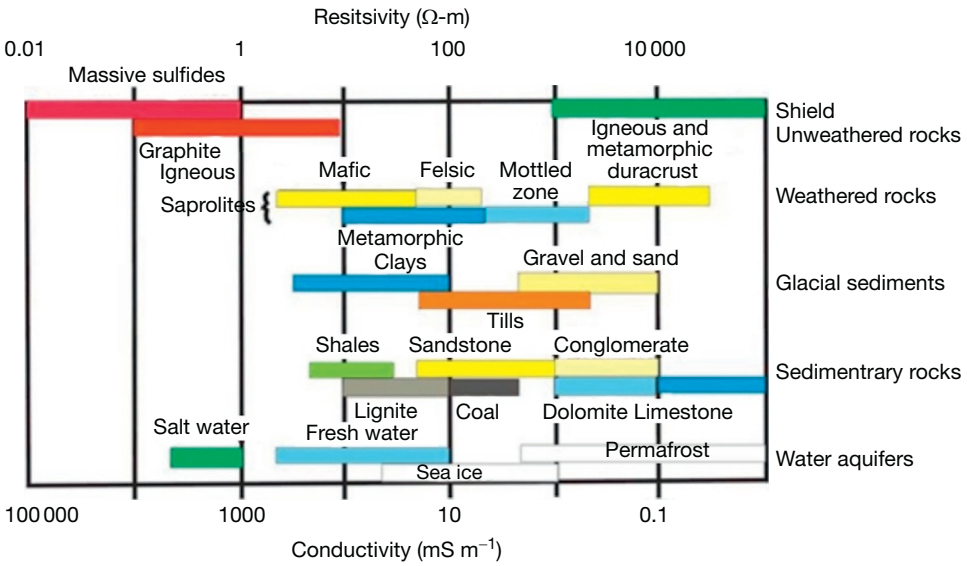


Figure 2 Resistivity (conductivity) ranges of rocks and sediments.

Table 5 Compressional (*P*-wave) velocity ranges of rocks and unconsolidated sediments

Type	*Velocity (m s^{-1})*
Unconsolidated sediments[a]	
Sand (dry)	200–1000
Sand (water saturated)	1500–2000
Clay	1000–2000
Glacial till (water saturated)	1500–2500
Sedimentary rocks[a]	
Sandstones	2000–6000
Tertiary sandstone	2000–2500
Carboniferous sandstone	4000–4500
Cambrian quartzite	5500–6000
Limestones	2000–6000
Cretaceous chalk	2000–2500
Jurassic oolite and bioclastic limestone	3000–4000
Carboniferous limestone	5000–5500
Dolomite	2500–6500
Salt	4500–5000
Anhydrite	4500–6500
Gypsum	2000–3500
Igneous/metamorphic rocks[a]	
Granite	5500–6000
Gabbro	6500–7000
Ultramafic rocks	7500–8500
Serpentinite	5500–6500
Fluids[a]	
Air	300
Water	1400–1500
Massive sulfide[b]	
Pyrite	7990
Sphalerite	5440
Chalcopyrite	5120
Pyrrhotite	4600
Pentlandite	4560
Bornite	4400
Galena	3690

[a]Modified from Kearey and Brooks (1984).
[b]Modified from Salisbury et al. (2003).

rocks range between 2000 and 6000 m s^{-1} and volcanic rocks tend to be >6000 m s^{-1}. Massive sulfide *P*-wave velocity is quite variable and ranges between 3700 and 7000 m s^{-1}. See Chapter 11.03 for additional information on seismic velocity.

11.15.3.5 Gamma Radiation from Rock and Minerals

The three naturally occurring radioelements are uranium (U), thorium (Th), and potassium (K). The concentration of uranium and thorium in rocks is measured in parts per million (ppm), and the concentration of potassium in rocks is measured in percent. Table 6 from Killeen (1979) provides these radioelement concentrations for different rock classes. The concentration is related to the number of radioactive decays (gamma radiation) per unit time usually measured in counts per second (cps). The more radioactive decays per unit time, the higher the concentration of radioelements. High-energy gamma rays are severely attenuated by rocks, unconsolidated sediments, and water. They generally only penetrate 30–50 cm before they are totally absorbed. Consequently, radiometric methods only measure radioactivity in the upper 30–50 cm of the Earth's surface. The radiometric technique was originally developed for uranium exploration since it directly measures uranium radioactivity. However, by using the concentrations of all three naturally occurring radioelements, it has evolved into an indirect geologic mapping technique. It is particularly effective at mapping potassic alteration (note that the mean concentration of K is 6.5% and 4.2% for alkali feldspathoidal intermediate extrusive and intrusive rocks, respectively). See Chapter 11.14 for additional information on gamma radiation.

11.15.3.6 Dielectric Constant of Rocks and Sediments

The velocity v of an EM wave in a material is given by $v = c/K^{1/2}$ where c is the velocity of light in free space (0.3 m ns^{-1}) and K is the dielectric constant of the material (Davis and Annan, 1989; Chapter 11.07). The dielectric constant of a material is

Table 6 Radioelement concentrations in different classes of rocks

	Potassium (%)		*Uranium (ppm)*		*Thorium (ppm)*	
Rock type	*Mean*	*Range*	*Mean*	*Range*	*Mean*	*Range*
Felsic extrusives	3.1	1.0–6.2	4.1	0.8–16.4	11.9	1.1–41.0
Felsic intrusives	3.4	0.1–7.6	4.5	0.1–30.0	25.7	0.1–253.1
Intermediate extrusives	1.1	1.1–2.5	1.1	0.2–2.6	2.4	0.4–6.4
Intermediate intrusives	2.1	0.1–6.2	3.2	0.1–23.4	12.2	0.4–106.0
Mafic extrusives	0.7	0.06–2.4	0.8	0.03–3.3	2.2	0.05–8.8
Mafic intrusives	0.8	0.01–2.6	0.8	0.01–5.7	2.3	0.03–15.0
Ultramafic	0.3	0–0.8	0.3	0–1.6	1.4	0–7.5
Alkali feldspathoidal Intermediate extrusives	6.5	2.0–9.0	29.7	1.9–62.0	133.9	9.5–265.0
Alkali feldspathoidal Intermediate intrusives	4.2	1.0–9.9	55.8	0.3–720.0 1	32.6	0.4–880.0
Alkali feldspathoidal Mafic extrusives	1.9	0.2–6.9	2.4	0.5–12.0	8.2	2.1–60.0
Alkali feldspathoidal Mafic intrusives	1.8	0.3–4.8	2.3	0.4–5.4	8.4	2.8–19.6
Chemical sedimentary rocks	0.6	0.02–8.4	3.6	0.03–26.7	14.9	0.03–132.0
Carbonates	0.3	0.01–3.5	2	0.03–18.0	1.3	0.03–10.8
Detrital sedimentary rocks	1.5	0.01–9.7	4.8	0.01–80.0	12.4	0.2–362.0
Metamorphosed igneous rock	2.5	0.1–6.1	4	0.1 – 148.5	14.8	0.1–104.2
Metamorphosed sediment rocks	2.1	0.01–5.3	3	0.1–53.4	12	0.1–91.4

Reproduced from Killeen PG (1979) Gamma-ray spectrometric methods in uranium exploration – Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in Search for Metallic Ores*. Geological Survey of Canada, Economic Geology Report 31, pp. 163–230.

defined as the ratio of dielectric permittivity of the material (ϵ) to the dielectric permittivity of free space (ϵ_o) where ϵ_o is equal to $8.85 \times 10^{-12}\,\mathrm{F\,m^{-1}}$, that is, $K = \epsilon/\epsilon_o$ (Reitz and Milford, 1962; Telford et al., 1976). Consequently, $K = 1$ in free space. The electric permittivity (ϵ) relates the electric field strength ($\mathbf{E}$) to the electric displacement ($\mathbf{D}$) through $\mathbf{D} = \epsilon\mathbf{E}$ and is a measure of how easy a material can be polarized when an external electric field is applied (Reitz and Milford, 1962). **Table 7** provides estimates of the dielectric constant for several rocks and materials of interest when carrying out GPR surveys (see **Chapter 11.07**).

Table 7 Approximate ranges of dielectric constant for different materials

Material	*Dielectric constant (K)*
Air (free space)	1
Water	80
Sedimentary rocks	
Shale	5–15
Limestone	4–8
Intrusive rocks	
Granite	4–6
Sediments	
Silts	5–30
Clays	5–40
Saturated sand	20–30
Dry sand	3–5

11.15.4 Geophysical Methods Applied to Mineral Exploration

Geophysical methods have been applied to mineral exploration for over a century (Ford et al., 2007). During the first half of the 1900s, magnetic, gravity, electrical, and refraction seismic methods were developed and applied to ground-based mineral exploration. In the late 1940s, EM techniques (e.g., slingram, one of the first horizontal loop EM systems) were developed and applied to ground-based mineral exploration. Airborne magnetic, EM, and radiometric techniques were also developed around this time. Since then, many significant advances have been made in geophysical instrumentation, acquisition, and processing. The application of modern computer techniques has greatly improved processing, displaying, and interpreting the data. The tools and techniques chapters in this book provide the basic theory and operating principles of the geophysical techniques. Consequently, this chapter focuses on their application for the direct detection of mineral deposits and for regional geologic mapping.

In order to have an observable response, all geophysical techniques require a contrast in physical properties of the target of interest and the surrounding rock. For example, there must be a density contrast between the target and the surrounding rock in order to see a gravity response. The physical properties of the target and surrounding rock therefore determine the geophysical technique or combination of techniques that will provide the best chance for detecting the target. In **Section 11.15.4.2**, the characteristics of individual deposit types and their associated host rocks (size, physical properties, geologic environment, etc.) are presented along with a discussion of which geophysical method or methods are most appropriate for their detection.

There are too many geophysical methods available to cover each of them in detail. However, they are described in detail in the earlier chapters within this volume. As an example, consider EM methods (Chapter 11.10). There are systems that operate in the frequency domain and others that operate in the time domain; there are airborne systems (fixed wing and helicopter) and ground systems; there are ground systems that use large loops; there are ground systems with different transmitter–receiver configurations; there are borehole systems; there are passive EM systems that use naturally occurring EM fields as a transmitter source (lightning and solar radiation); there are also EM systems that are designed to explore deep; and there are other systems that are designed to provide detailed information of the top 100 m or less of the subsurface.

Other geophysical methods, although not as complex as EM methods, still have a multitude of different hardware choices to consider. In addition, there are many different schemes for processing, presenting, and interpreting geophysical data. One-dimensional (1-D), 2-D, and 3-D inversion programs have been developed for almost all the geophysical methods, as well as sophisticated programs for gridding and imaging the data. Therefore, this chapter focuses on how the different geophysical techniques and their presentation and interpretation schemes are used (1) for geologic/regional mapping and (2) for direct detection of specific mineral deposit types.

11.15.4.1 Geophysical Applications to Regional/Geologic Mapping

Regional geophysical methods can provide important information on geologic structure, lithologic information, and alteration patterns that can be used to narrow down the search for specific mineral deposits and for general geologic mapping. Modern regional geophysical surveys are generally carried out using airborne platforms, although ground-based geophysical surveys provide regional information in specific cases. Examples of regional magnetic, gravity, and radiometric maps are provided in the succeeding text. Further examples of utilizing regional geophysical surveys to focus exploration activity for individual mineral deposit types will be presented later where appropriate.

Magnetic surveys are routinely used for mapping geology at a regional scale. Airborne magnetic and magnetic gradient surveys are often carried out as a first step in an exploration program (Chapter 11.11). Many techniques have been developed to process, display, and interpret magnetic data qualitatively. Quantitative techniques have been developed as well, for example, 2-D and 3-D Euler depth estimation, which provides quantitative information on the depth to magnetic sources. Quantitative estimates of magnetic susceptibility using 2-D and 3-D inversion methods are used to identify different rock units (Table 3). Figure 3 is an example of a 3-D magnetic susceptibility map computed at sea level using the total magnetic field data from the QUEST Project (Geoscience BC) situated in central British Columbia. The area is covered with thick glacial sediments, and since the surface elevation is several hundred meters above sea level, the sea-level susceptibility map is imaging bedrock. It was computed by Mira Geoscience for Geoscience BC's QUEST Project using the University of British Columbia Geophysical Inversion Facility magnetic 3-D susceptibility inversion software (Li and Oldenburg, 1998). The blue colors are low susceptibility values representing rocks with little or no magnetite in them, while the yellow to orange colors represent rocks containing more magnetite (mafic and ultramafic rocks).

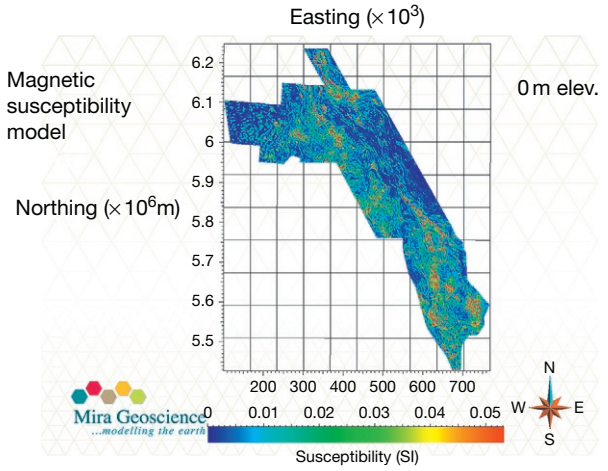


Figure 3 Example of using 3-D inversion to generate a magnetic susceptibility map from the total magnetic field. The magnetic susceptibility varies between 0 (blue) and 0.05 (red) SI units. Reproduced from Geoscience BC's QUEST Project.

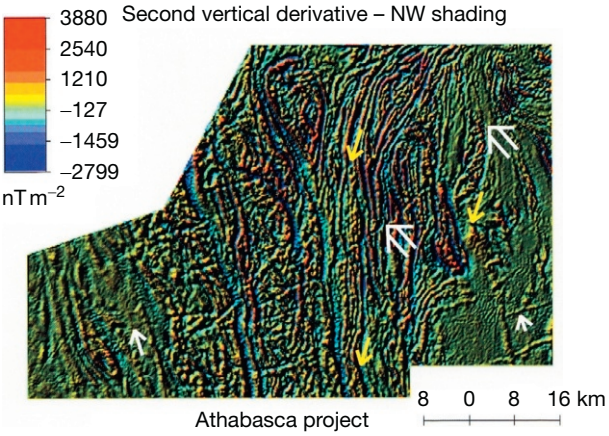


Figure 4 Second vertical derivative map of the total magnetic field with sun shading from the northwest. Reproduced from Best ME, Abercrombie HJ, and Peirce JW (1998) Interpreted faulting patterns in northeast Alberta using high-resolution aeromagnetic data. *Canadian Journal of Exploration Geophysics* 34: 48–57.

Magnetic survey data can also be manipulated by applying different filters and derivatives to the data to enhance the location of faults and lineaments. This provides qualitative information on geologic structure and lithologic boundaries. Figure 4 is a map of the second vertical derivative of the total magnetic field with sun shading from the northwest (Best et al., 1998). The magnetic data were collected for a sedimentary gold and kimberlite project near Fort McMurray in NE Alberta. The white arrows point to the location of several WSW–ENE lineaments, and the double white arrows point to the location of a major NE–SW lineament. This lineament is offset in a left lateral sense by a NW–SE fault just to the SW of the center white double arrow.

Gravity surveys provide regional information as well (Chapter 11.12). Ground gravity surveys, which until recently were the only surveys available, are time-consuming and costly, so they usually consisted of sparse data sets. However, modern airborne gravity surveys can quickly cover large areas with dense coverage, thus increasing the usefulness of the data for more detailed interpretation. Many of the techniques used for qualitatively interpreting magnetic data (filtering, derivative maps, upward/downward continuation, and sun shading) are also applied to gravity data. Quantitative estimates of density using 2-D and 3-D inversion methods are also used to identify different rock units (Table 4). Figure 5 is an example of a 3-D density contrast map computed at sea level from airborne Bouguer gravity data for the same Geoscience BC's QUEST Project. Mira Geoscience computed the 3-D density from the Bouguer gravity data using the University of British Columbia Geophysical Inversion Facility 3-D density inversion software (Li and Olderburg, 1998). Regions with larger positive density contrasts (red) tend to line up with the higher magnetic susceptibility regions seen in Figure 3 as expected since mafic rocks are usually dense as well as being magnetic.

Radiometric methods (see Chapter 11.14) measure the decay of the natural occurring radioelements (U, Th, and K). The higher the concentration of one of these three radioelements, the more decays occur per second. For example, a higher potassium count rate in a given area indicates that the area contains a higher percentage of potassium. Such information can be used to map certain types of alteration in rocks and minerals that are associated with geologic processes. The measurement of uranium and thorium count rates provides a method for the direct detection of uranium mineralization. Radiometric techniques, however, can only map the upper 30 cm or so of the Earth's surface since the high-energy gamma rays emitted are absorbed in a very short distance by rock, water, and unconsolidated sediments. Consequently, the radiometric technique is best suited to areas with little or no cover. Figure 6 is an example of a map of potassium and thorium counts per second from a regional radiometric survey in Africa (Ruotoistenmäki, 2008). Maps like this can be used to outline geologic formations and potential alteration regions. The red colors indicate areas that contain higher concentrations of potassium and thorium, and the white areas are covered by wetlands and lakes. Note that the Archean and Proterozoic rocks in the northwest (consisting of igneous and intrusive rocks) are highly anomalous compared with the rest of the area.

11.15.4.2 Geophysical Applications by (Metal) Deposit Type

Tectonics is the study of the movements and deformation of the Earth's crust on a large scale. Plate tectonics describes the

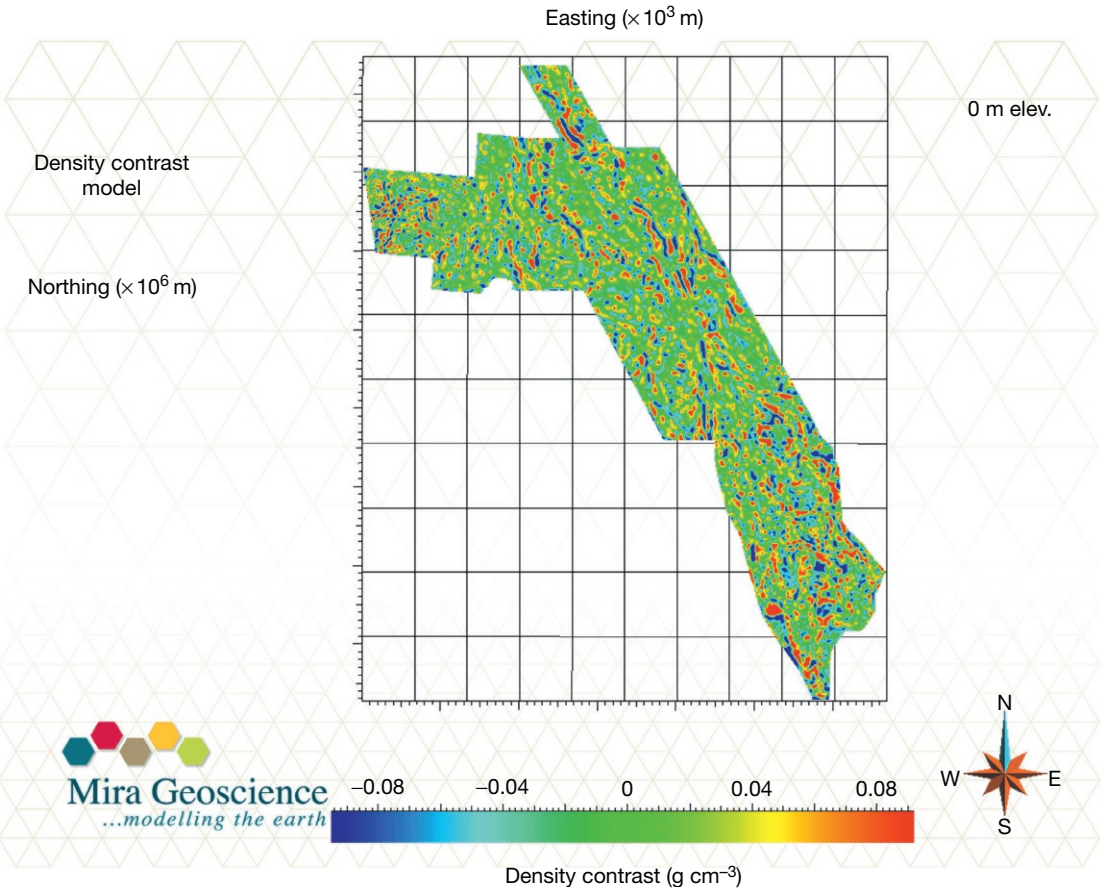


Figure 5 Example of using inversion to generate a density contrast map from the Bouguer gravity data. The density contrast varies between $-0.08\ g\,cm^{-3}$ (blue) and +0.08 (red). Reproduced from Geoscience BC's QUEST Project.

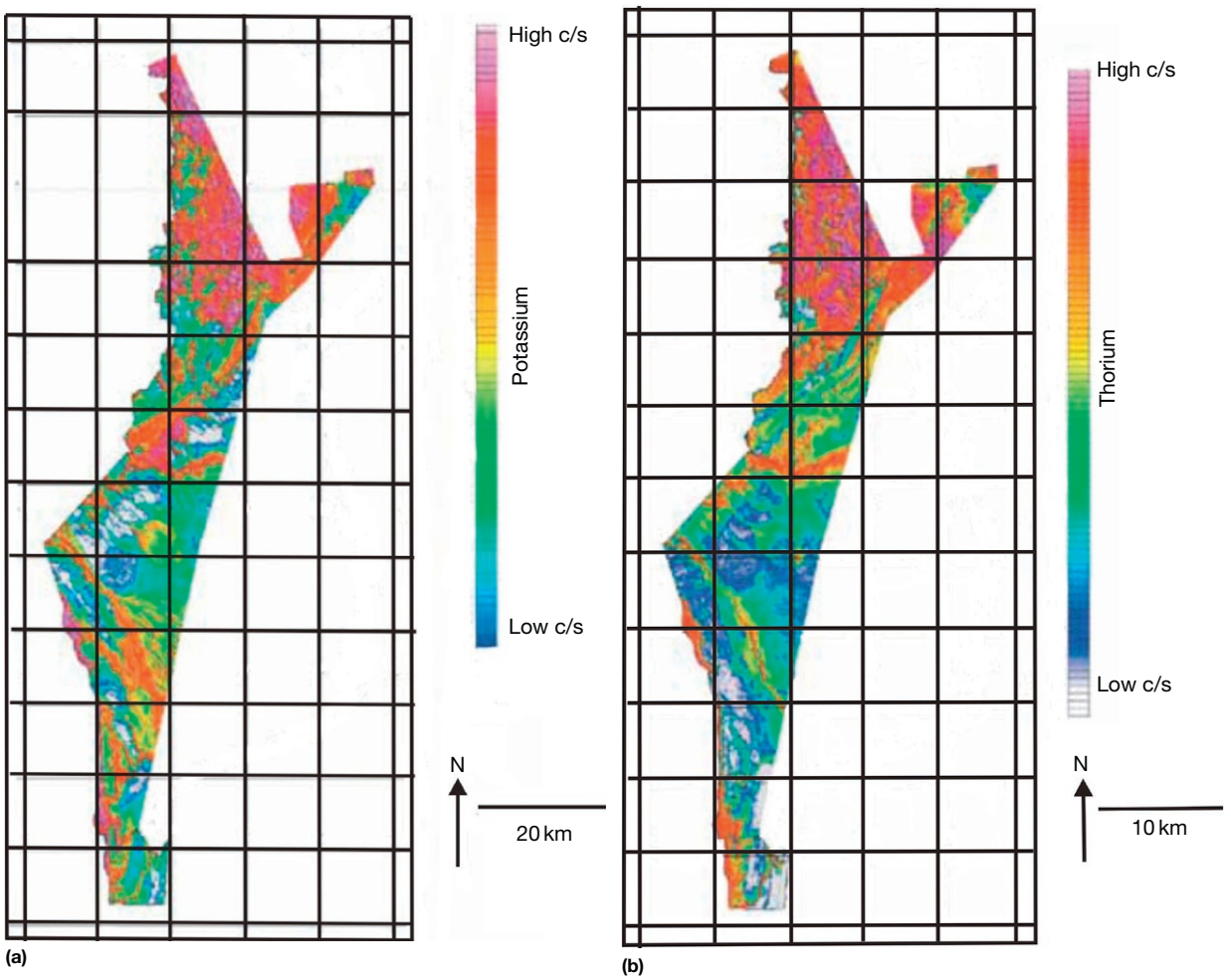


Figure 6 Example showing maps of (a) potassium counts per second c/s and (b) thorium counts per second from an airborne survey in Mozambique. Modified from Ruotoistenmäki T (2008) Geophysical maps and petrophysical data of Mozambique. *Geological Survey of Finland*, *Special Paper 48*, pp. 65–80.

large-scale motions of the Earth's lithosphere, which is made up of plates or zones, for example, the North American and Pacific Plates (Meissner, 2002; Wegener, 1966). There are three types of plate boundaries: (1) Convergent boundaries occur where two plates slide toward each other, (2) divergent boundaries occur where two plates slide apart from each other, and (3) transform boundaries occur where two plates slide past each other. Many mineral deposits are associated with plate boundaries because these boundaries allow magma and hydrothermal fluids deep within the Earth's crust to travel up along the boundary, eventually reaching the shallow subsurface. Examples of the mineral deposits associated with these boundaries will be presented in this section.

There are so many different types of mineral deposits that all of them are not covered in detail in this chapter. The focus is on several deposit types that generate significant volumes of economic minerals: (1) VMS deposits, (2) porphyry deposits, (3) kimberlites (diamonds), (4) unconformity-related uranium deposits, (5) laterite (nickel) deposits, and (6) placer (gold, diamonds, tin, and aggregate) deposits. In addition, a few comments will be made concerning sedimentary exhalative sulfide (SedEx) deposits (Pb–Zn and Cu), hydrothermal gold deposits, magmatic Ni–Cu–PGE's deposits, and carbonate Ag–Pb–Zn replacement deposits.

A short description of the geologic characteristics of these deposits is provided, along with references for those wishing more detailed information. The geophysical techniques appropriate for each of the deposit types mentioned in the preceding text are discussed in this section. Two good review articles have been published in the last decade discussing recent geophysical developments for mineral exploration (Nabighian and Asten, 2002; Vallee et al., 2011). They have many useful references containing more detailed information on geophysical applications for mineral exploration, including new developments in display, presentation, modeling, and inversion of geophysical data.

11.15.4.2.1 VMS deposits

VMS deposits form at or near the seafloor in submarine volcanic environments containing metal-enriched fluids associated with seafloor hydrothermal convection related to spreading centers. The deposits usually have two parts; the massive section is typically a mound-shaped stratabound body consisting principally of massive (>40%) sulfide, quartz, and iron oxide minerals and altered silicate wall rock. These stratabound

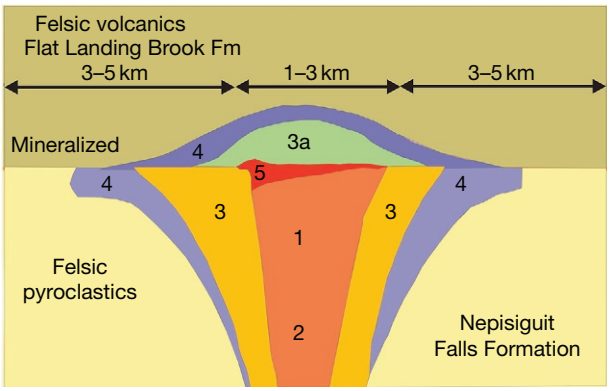


Figure 7 Composite cross section of a VMS massive sulfide deposit showing alteration zones from the Bathurst mining camp, Canada. Zone 1, quartz plus Fe chlorite; Zone 2, Fe chlorite plus sericite; Zone 3, Fe–Mg chlorite plus sericite; Zone 3a, phengite plus chlorite; Zone 4, phengite plus Mg chlorite; Zone 5, sulfide zone. Modified from Goodfellow WD, McCutcheon SR, and Peter JM (2003) Massive sulfide deposits of the Bathurst Mining Camp, New Brunswick and Northern Maine: Introduction and summary of findings. *Economic Geology Monograph* 11: 1–16.

bodies are often underlain by stockwork veins and disseminated sulfides (Barrie and Hannington, 1999a,b; Franklin et al., 1981, 2005; Lydon, 1984). The stockwork vein systems are enveloped in distinctive alteration halos (**Figure 7**), which may extend into the hanging-wall strata above the VMS deposit (Goodfellow et al., 2003; Lydon, 1984).

VMS deposits are major sources of the base metals Cu, Zn, Pb, as well as Au and Ag. They are classified according to metal content, gold content, and host-rock lithology and are divided into Cu–Zn, Zn–Cu, and Zn–Pb–Cu groups based on the ratio of Cu, Zn, and Pb contained in them (Franklin, 1996). The deposits contain the sulfide minerals chalcopyrite, sphalerite, galena, pyrrhotite, and pyrite as well as magnetite. They occur in submarine volcanic terranes associated with seafloor spreading and range in age from Archean to present day (such as the Juan de Fuca Ridge located several hundred kilometers off the coast of Western Canada and the United States (Davis et al., 1987; Duckworth et al., 1994)). These deposits typically occur in clusters that define mining camps, for example, the Bathurst Mining Camp in Eastern Canada.

Several useful review papers on geophysical methods for VMS exploration have been previously published (Boivin, 2007; Morgan, 2012) and a paper on geophysical methods applied to several deposit types by Ford et al. (2007). These provide additional information and references.

Pyrrhotite mineralization is often associated with VMS deposits. Since pyrrhotite is a magnetic mineral (**Table 3**), there can be a magnetic anomaly associated with VMS deposits containing pyrrhotite. On the other hand, regional magnetic surveys are only moderately effective for VMS exploration. The massive section of VMS deposits will be a good conductor as long as the sulfide grains are in contact with each other to form a continuous path for electric current to flow (Keller, 1988; Palacky, 1988). Although sphalerite is not a good conductor, it is always associated either with chalcopyrite or galena, which are good conductors (Shuey, 1975). The volcanic and metasedimentary rocks surrounding VMS deposits are resistive (**Figure 7**). The contrast in conductivity between the massive sulfide body and the surrounding rocks therefore can be more than 100:1 making them ideal EM targets. Indeed, ground EM methods were originally developed to locate massive sulfide ore bodies. However, not all massive sulfide targets are easy to detect (Cooper and Swift, 1994). Airborne EM methods were developed in the late 1940s to locate massive sulfide ore bodies (Collett, 1986). Ground and borehole EM techniques are also effective for mapping VMS deposits (Boivin, 2007; Cooper and Swift, 1994; Hattula and Rekola, 2000).

Resistivity methods are applied on the ground because they require direct contact with the Earth (see chapter on electrical methods in **Chapter 11.08**). The same comments about conductivity contrasts apply to these techniques as well although it is more common to talk about resistivity for electrical methods rather than conductivity. Electrical methods are therefore quite effective for the direct detection of VMS deposits. The sulfide minerals in the massive section of VMS deposits have higher densities (**Table 4**) than the surrounding rocks. The density contrast, however, is much smaller than the conductivity contrast, typically 1.3:1 or less. Radiometric methods can be effective for mapping regional alteration associated with VMS targets, but they are only moderately effective for more detailed target identification. Potassic alteration in the form of sericite (**Figure 7**) is commonly associated with many types of VMS base-metal and gold deposits (Franklin, 1996; Poulsen and Hannington, 1996; Shives et al., 2000).

Figure 8 is a geologic cross section of the Caber deposit in the Matagami VMS mining camp of western Quebec, Canada. The deposit is sphalerite-rich, cigar-shaped, ~30 m wide by 250 m in strike length, steeply SW-dipping, and buried between 120 and 350 m, below ~30 m of conductive clay overburden. It is a difficult target for EM since it is small and not very conductive and occurs beneath conductive overburden.

Geotech Airborne Geophysical Surveys of Aurora, Ontario, Canada, flew the Caber deposit in 2012 using their 35 m transmitter loop (dipole moment of ~1.3 million Am^2) helicopter VTEM system. **Figure 9(a)** is a plan view of the flight lines over the deposit showing tau values (Yang and Oldenburg, 2012). The tau value is the decay constant of an exponential decay curve fit to $d\mathbf{B}/dt$ values for a specific set of voltage channels where $d\mathbf{B}/dt$ is the time rate of change of the magnetic field. The decay channels selected determine the approximate depth of the tau function. Higher tau values are caused by more conductive material in the subsurface at approximately that depth range. **Figure 9(b)** shows the resistivity depth image (Fullagar, 1989; Huang and Rudd, 2006; Macnae et al., 1991) for one of the flight lines on the bottom and the VTEM $d\mathbf{B}/dt$ channels along the top. Note that the deposit appears as a conductivity high at depth.

Figure 10 is a plot of the $d\mathbf{B}/dt$ decay channels for the z (top) and x (bottom) receivers with the resistivity depth image for line 1040 at the bottom. The nearly vertical blue dike in the resistivity depth image marks the position of a plate model used to compute the $d\mathbf{B}/dt$ decay channel responses. The black curves on the decay channels are the measured $d\mathbf{B}/dt$ values, and the red curves on the decay channels are the computed $d\mathbf{B}/dt$ values from the plate model. Note that the top of the plate model matches the position of the top of the conductivity

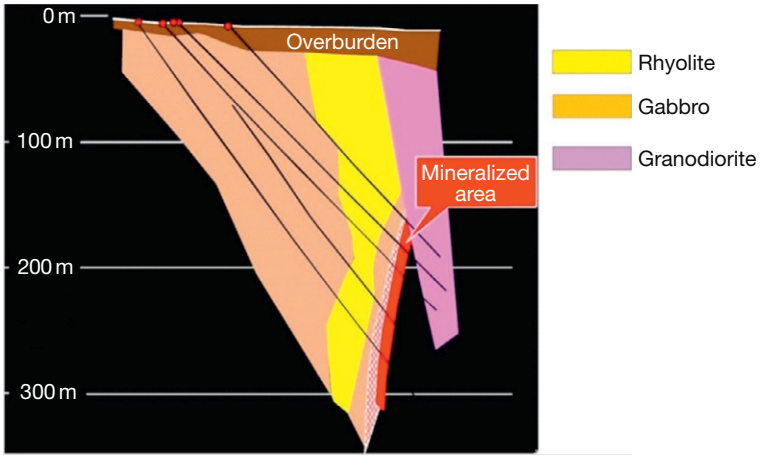


Figure 8 Geologic cross section of the Caber deposit. The mineralized area outlined on the figure is the massive sulfide ore body. Courtesy of Geotech airborne Geophysical Surveys.

anomaly in the resistivity depth image that is associated with the deposit. This example illustrates how modern airborne EM methods can detect small, relatively low-conductivity VMS deposits at depths >150 m.

11.15.4.2.2 Porphyry deposits

Porphyry deposits are large, low- to medium-grade deposits in which primary (hypogene) ore minerals are dominantly structurally controlled and are spatially and genetically related to felsic to intermediate porphyritic intrusions (Kirkham, 1971, 1972). The main metals in a porphyry are Cu, Mo, and smaller amounts of gold, although Sn, W, and Ag can be found in some porphyry deposits (Kirkham and Sinclair, 1996; Sinclair, 2007). Deposit size ranges from tens of millions to billions of tons; grades for the different metals vary considerably but generally average <1% (Sinclair, 2007). They range in age from Archean to the present, although the majority of economic deposits are Jurassic or younger (Sinclair, 2007). **Figure 11** is a schematic representation of the tectonic setting of a porphyry deposit, and **Figure 12** illustrates typical zoning in a Cu porphyry deposit. Minerals of importance for geophysical detection include chalcopyrite, bornite, pyrite, magnetite, and alteration products (e.g., biotite, K-feldspar, and sericite).

Magnetic surveys can be used to outline porphyry Cu and Cu–Au deposits, which have abundant hydrothermal magnetite-, pyrrhotite-, and/or magnetite-bearing skarn and hornfels zones around porphyry-related intrusive rocks. Conversely, some deposits are characterized by magnetic lows due to the destruction of magnetite in phyllic alteration zones. EM methods can be used to map the depth and lateral extent of porphyry systems. For example, controlled source audio-frequency magnetotelluric (CSAMT) surveys have proven to be effective for defining the horizontal limits and depth extent of porphyry systems (Corry et al., 1988). Ground and airborne time-domain EM methods have also been used to map porphyry systems (Pare and Legault, 2010; Xue et al., 2012; Yang and Oldenburg, 2012).

The dispersed nature of sulfide distribution in porphyry deposits is amenable to the induced polarization (IP) method, which was originally developed for porphyry Cu exploration (Brant, 1966; Gasperikova et al., 2005; Pare and Legault, 2010; Seigel et al., 2007). Gamma-ray spectrometry surveys have been used to outline potassic alteration zones closely related to mineralized zones, for example, the Mount Milligan deposit in central British Columbia and the Casino deposit in west-central Yukon Territory (Shives et al., 2000). Gravity methods have only been marginally effective for mapping porphyry systems.

The following example from Pare and Legault (2010) illustrates some of the subtleties when interpreting geophysical data over a porphyry system. The Pebble deposit in Alaska, one of the largest Cu–Au deposits in the world, is a calc-alkalic porphyry deposit. Mineralization at Pebble West occurs around small granodioritic stocks that intrude the country rocks. The Pebble East mineralization occurs within a granodioritic stock and in sills that cut the country rocks. The Pebble West zone outcrops (**Figure 13**) but the Pebble East zone is beneath thick cover up to 600 m thick (Rebagliati et al., 2008). The deposit is bounded in the SE by a major fault (labeled the ZGI fault in **Figure 13**).

The deposit hosts K-silicate alteration and associated quartz-sulfide veins, overprinted by phyllosilicate alteration. Sulfides mainly consist of hypogene pyrite, chalcopyrite, molybdenite, and bornite; thin oxide zones also occur at Pebble West.

Several 2-D dipole–dipole and pole–dipole IP/resistivity surveys were conducted within the area. The 'a' spacing (150 and 500 m) used depended on the specific data set; however, n values up to eight were used in all data sets. The 'a' spacing is the distance between the two current (voltage) electrodes, and the 'n' value represents the separation 'n multiplied by a' between the current and voltage electrodes (Sumner, 1976). Spacing between the 2-D lines was no more than 500 m. **Figure 14(a)** is a depth image (100–200 m) of the 2-D IP inversion. There is a belt of high chargeability (red colors) surrounding Pebble West that extends from ~2 km north of the deposit to the bottom SW corner of the survey area. There is no direct correlation between the high IP response and the economic mineralization at Pebble West; the chargeability is mostly mapping pyrite and alteration zones around the deposit. There is a chargeability low around Pebble East.

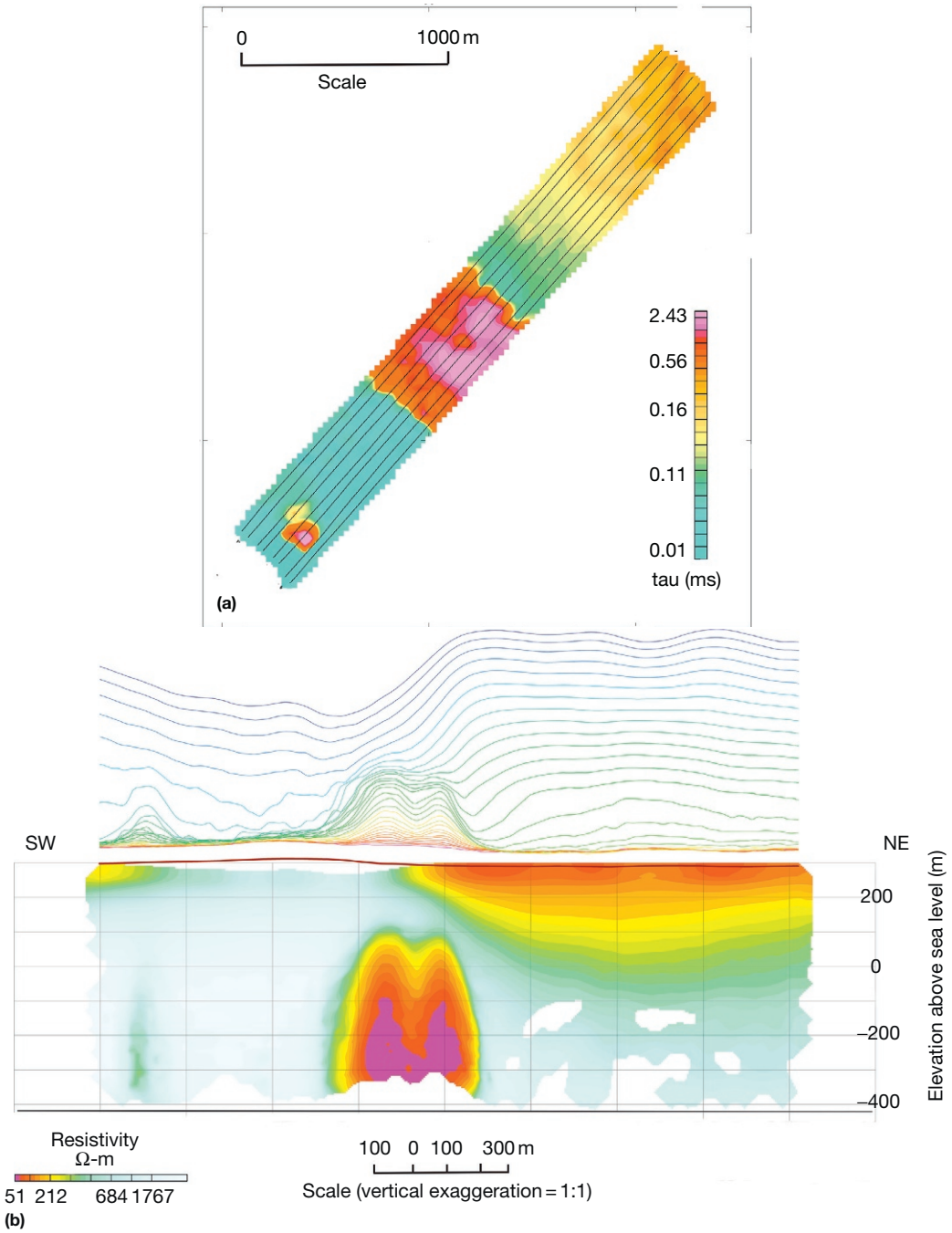


Figure 9 Geotech VTEM^{MAX} airborne EM results over the Caber deposit. (a) Plan view of d**B**/dt tau values with high tau values centered over the Caber deposit. (b) Resistivity depth image cross section with d**B**/dt channels along the top (increasing time downward) for line 1040. The high-conductivity (low-resistivity) feature at depth is associated with the deposit body. Courtesy of Geotech airborne Geophysical Surveys.

The resistivity values are generally lower north of the deposits than south of the deposits (Figure 14(b)). The Pebble West deposit does not have a unique resistivity signature, so the resistivity data do not directly locate the economic mineralization. The Pebble East deposit was not observed on these data sets because the survey parameters were not designed to penetrate to that depth. This survey shows how difficult it can be to directly locate economic mineralization within a porphyry system. Often, these techniques can only provide a broader indication of alteration patterns associated with these porphyry systems. However, they do provide guidance where to focus on further exploration activity.

Alteration associated with porphyry deposits can be mapped quiet effectively using radiometric methods. The following example from the paper by Shives et al. (2000) illustrates how radiometric methods can be used to locate porphyry

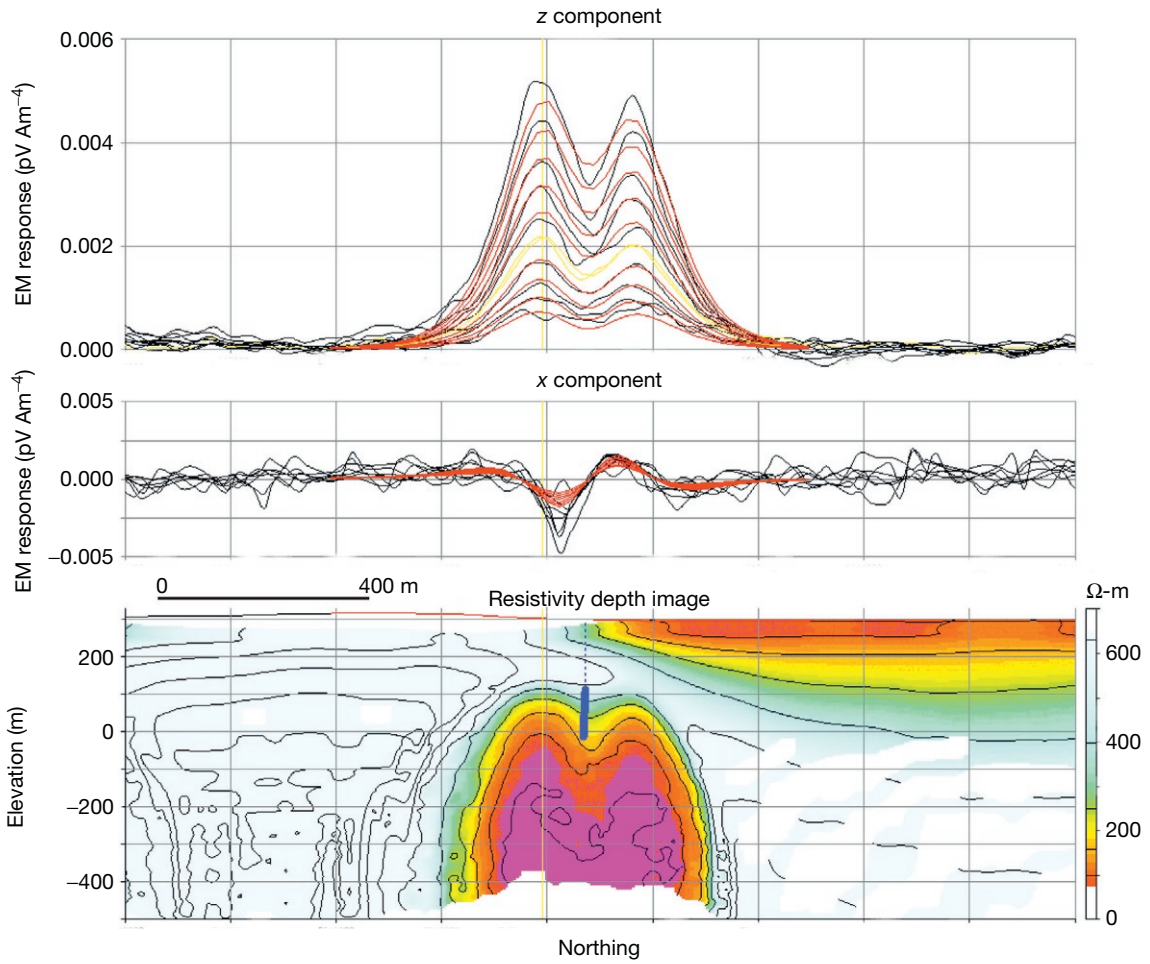


Figure 10 The upper two plots are the d**B**/d*t* decay channels of the *z* (top) and *x* (bottom) receivers for line 1040. The black curves are the measured d**B**/d*t* values, and the red curves are the d**B**/d*t* values computed from the plate model shown as the nearly vertical blue line in the center of the bottom resistivity depth image. Courtesy of Geotech airborne Geophysical Surveys.

alteration. The Casino Au–Cu–Mo porphyry deposit is located in Northern Canada in deeply weathered, unglaciated terrain. The deposit is associated with the upper Cretaceous Casino Intrusive Complex, which intrudes middle-Cretaceous granodiorites. A preserved, 70 m thick, gold-bearing, leached cap overlies a Cu-rich supergene zone, with the base of weathering extending to 300 m below the surface. A central, mineralized, potassic alteration zone consisting of K-feldspar and biotite is surrounded by barren phyllic (sericite) and propylitic alteration (Bower et al., 1995; Shives et al., 2000).

Figure 15 shows that despite deep weathering, the in situ spectrometer radioelement concentrations can differentiate host lithologies and alteration. Figure 16 shows maps of K concentration and eTh/K ratio from an airborne radiometric survey with 500 m line spacing. Potassium is determined directly from the measured gamma rays, but thorium and uranium are determined indirectly by counting the gamma rays from daughter products in the decay series, thallium (Tl) 208 and bismuth (Bi) 214, and assuming 'radioactive equilibrium' in the decay of thorium and uranium. Thorium and uranium can then be determined by counting gamma rays associated with the daughter products Tl and Bi. The 'e' means 'equivalent' because Th and U are not measured directly but are determined from their 'equivalent' amount that would be present to explain the Tl and Bi counts calculated.

The low eTh/K ratio anomaly in Figure 16(b) distinguishes high potassium values associated with altered, mineralized bedrock in the Casino deposit from similar high-potassium patterns (Figure 16(a)) associated with felsic volcanic rocks in the surrounding Yukon Metamorphic Terrane, which has higher eTh/K values. The stacked profiles in Figure 16(c) illustrate the dramatic drop in eTh concentration over the Casino deposit.

11.15.4.2.3 Kimberlite (diamonds) deposits

Kimberlites are formed at depths > 150 km beneath the Earth's surface. They have an igneous-rock matrix composed of carbonate, garnet, olivine, pyroxene, and upper mantle rock with several unique trace minerals (Bergman, 1987). It is thought that they erupt in a short time and travel to the Earth's surface rapidly. As they approach the Earth's surface, they form pipe-like features (Figure 17) – semicircular in shape (Kjarsgaard, 2007; Mitchell, 1986). The surface of the kimberlite can weather over time leading to lakes forming in the craters.

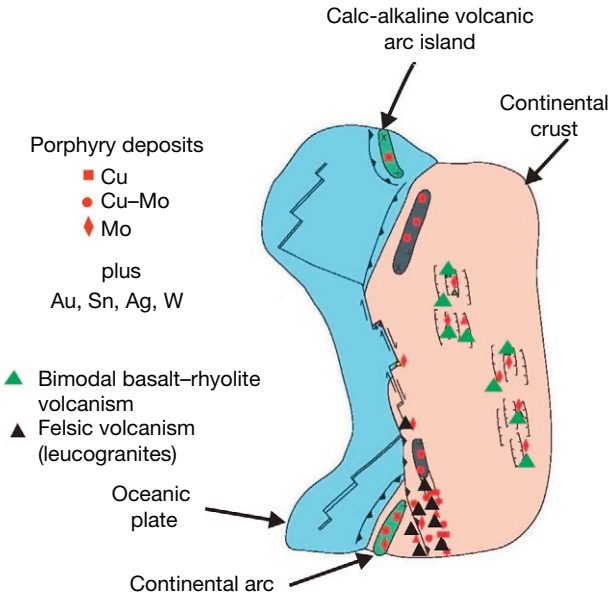


Figure 11 Schematic diagram showing the tectonic setting of porphyry deposits. The red dots indicate potential locations of different types of porphyry deposits relative to the boundary between the oceanic crust and continental crust. Modified from Kirkham RV and Sinclair WD (1995) Porphyry copper, gold, molybdenum, tungsten, tin, silver. In: Eckstrand OR, Sinclair WD, and Thorpe RI (eds.) *Geology of Canadian Mineral Deposit Types*, pp. 421–446. Geology of Canada. Geological Survey of Canada, No. 8.

Diamonds can be found in these pipes, although most kimberlite pipes do not contain diamonds and even fewer contain gem-quality diamonds. The surface expression of kimberlite pipes is small, typically a few hundred meters in diameter. Diamonds cannot be detected directly using geophysical methods. However, the geophysical signature of the kimberlite pipes hosting the diamonds is often different from the surrounding host rocks (Macnae, 1979).

Kimberlites often contain magnetite, which makes magnetic methods effective for locating these small targets (Jansen and Witherly, 2004; Keating and Sailhac, 2004; Morris et al., 2002). EM methods can be effective if the weathered zone at the surface is more conductive than the surrounding surface material (Jansen and Witherly, 2004; Munday et al., 2004; Smith et al., 1996). Similar comments apply to resistivity methods. Kimberlites are generally less dense than the surrounding rocks making them relatively easy to locate using gravity methods (Foss, 2004; Vasanthi and Mallick, 2005).

The Tli Kwi Cho kimberlite complex is located ~360 km NE of Yellowknife, Northern Canada (**Figure 18(a)**) in the Lac de Gras kimberlite cluster of the Archean Slave craton (Jansen and Witherly, 2004). The complex was found during a helicopter magnetic and EM survey flown in 1992. **Figure 18(b)** is a map of the total magnetic intensity from the helicopter survey showing the location of the two Tli Kwi Cho anomalies. The line spacing was 200 m, and the magnetic and EM sensor heights were at 50 and 20 m, respectively. The magnitude of the D027 anomaly is ~240 nT, with a resistivity low slightly

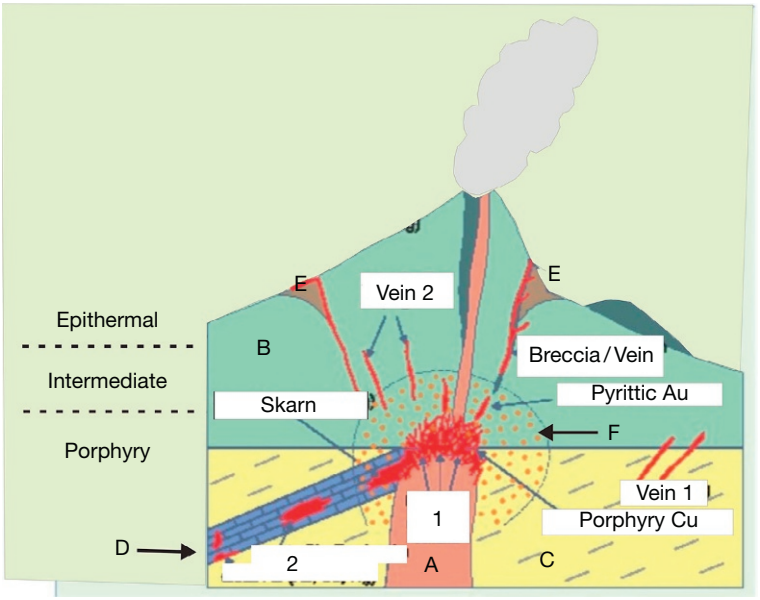


Figure 12 Schematic cross section of a Cu porphyry system in the root zone of an andesitic stratovolcano showing mineral zonation and possible relationship to skarn (Cu, Au, Mo, and Ag), manto (Pb, Zn, Ag, and Au), 'mesothermal' or 'intermediate' precious-metal and base-metal vein (vein1 = Au, Ag; vein2 = Pb, Zn, Au, Ag) and replacement (breccia/vein = Au, Ag), and epithermal precious-metal deposits, A = porphyry intrusion; B = volcanic rocks; C = clastic sedimentary rocks; D = carbonate-rich sedimentary rocks; E = Argillic alteration; F = pyritic zone. Modified from Kirkham RV and Sinclair WD (1995) Porphyry copper, gold, molybdenum, tungsten, tin, silver. In: Eckstrand OR, Sinclair WD, and Thorpe RI (eds.) *Geology of Canadian Mineral Deposit Types*, pp. 421–446. Geology of Canada. Geological Survey of Canada, No. 8.

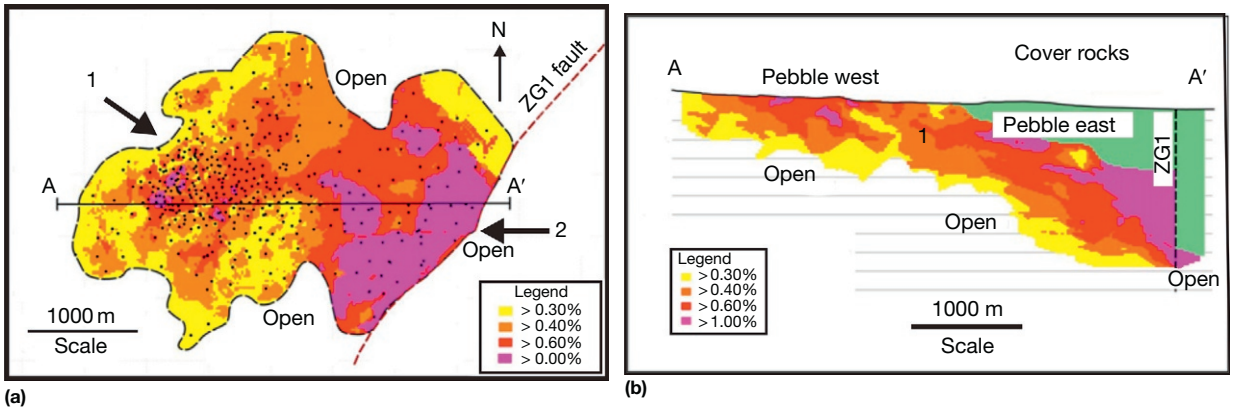


Figure 13 (a) Plan view of Pebble deposit: (1) Jura–Cretaceous to Eocene igneous and sedimentary rocks and (2) Late Cretaceous to Early Tertiary volcanic and sedimentary rocks. The units for the legend are copper equivalent percent. (b) Cross section of Pebble deposit showing general geology (1 represents same geology as in (a)) and grade distribution (legend same as (a)) across Pebble West and Pebble East zones. Vertical to horizontal scale 1:1. Modified from Pare P and Legault JM (2010) Ground IP-resistivity, and airborne spectrem and helicopter ZTEM survey results over Pebble copper-moly-gold porphyry deposit, Alaska. *Society of Exploration Geophysicists Annual Meeting, Denver, Colorado*.

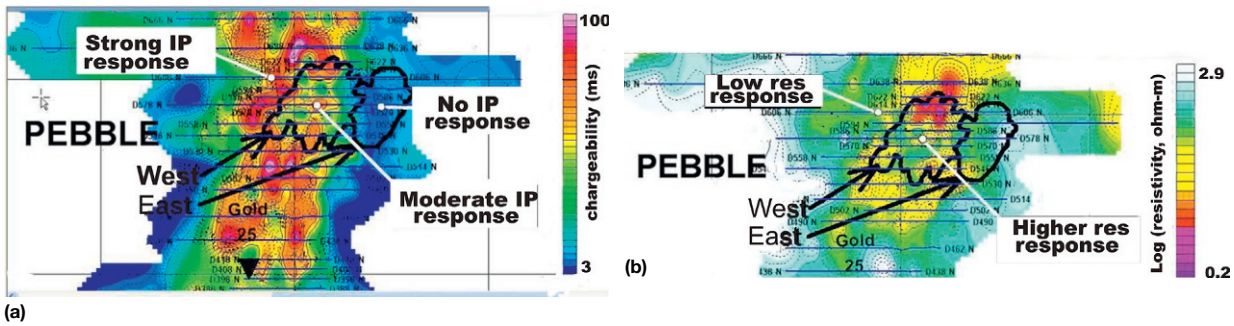


Figure 14 Level plan (100–200 m elevation) showing (a) chargeability (high values in red) and (b) resistivity (low values in red and purple) over the Pebble deposit from 2-D inversion. Modified from Pare P and Legault JM (2010) Ground IP-resistivity, and airborne spectrem and helicopter ZTEM survey results over Pebble copper-moly-gold porphyry deposit, Alaska. *Society of Exploration Geophysicists Annual Meeting, Denver, Colorado*.

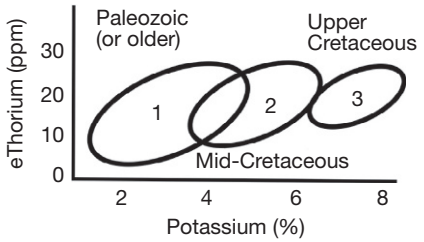


Figure 15 Potassium versus thorium concentrations from ground spectrometry within the vicinity of the Casino deposit: (1) Yukon Metamorphic Complex, gneiss, amphibolite, and quartzite; (2) Dawson Batholith, quartz diorite and granodiorite; (3) Casino intrusions, mineralized felsic microbreccia. Modified from Shives RBK, Charbonneau BE, and Ford KL (2000) The detection of potassic alteration by gamma-ray spectrometry—Recognition of alteration related to mineralization. *Geophysics* 65: 2001–2011.

offset from the magnetic high. Both these anomalies were located partly under a lake.

Zonge Engineering and Research of Tucson, AZ, carried out a time-domain NanoTEM survey over the area with a 40 m station spacing and a 50 m line separation. The system was operated in the central sounding mode. It is designed to measure early time gates after the transmitter current is turned off in order to see weakly conductive, shallow anomalies. Figure 19 shows the results of the survey for two time channels (channel 4 at early time and channel 10 at later time). The D018 and D027 zones are clearly outlined in the data.

A gravity low is expected over the kimberlite pipes due to the lower density of kimberlite ($\rho_{avg} = 2200$ kg m^{-3}) with respect to the host rocks ($\rho_{avg} = 2670$ kg m^{-3} for granite). Figure 20(a) is the Bouguer gravity from a ground gravity survey over the two pipes. Note that the gravity data suggest that the pipes may be connected. Neither the magnetic nor EM data suggested such a connection. A 2.5-D model (3-D volume with constant physical properties) of the gravity was calculated (Figure 20(b)). The calculated gravity (Figure 20(b)) agrees reasonably well with the Bouguer gravity in Figure 19(a). The depth of the pipe extended to 250 m as confirmed through drilling. However, deeper kimberlite could exist. Figure 20(c) is the forward gravity model for D018, assuming a density contrast of −300 kg m^{-3}.

11.15.4.2.4 Unconformity-related uranium deposits

Unconformity-related uranium deposits are located at or near the base of sandstone basins, which overlay deformed

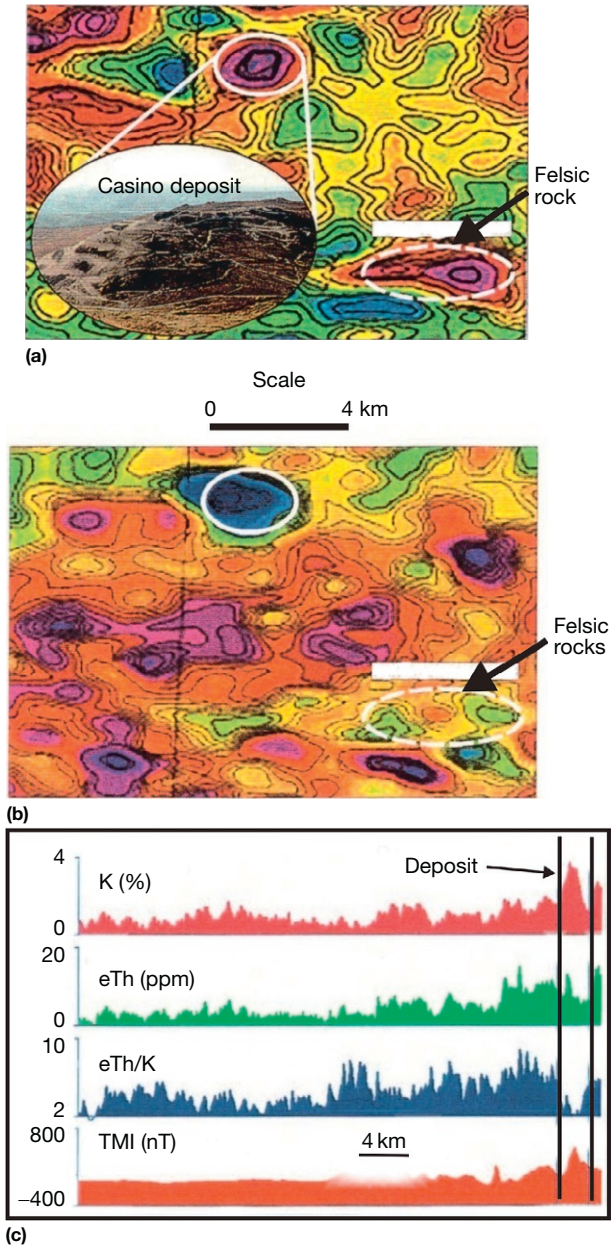


Figure 16 (a) Potassium concentration (red = high K %), (b) eTh/K ratio (blue = low ratio), and (c) potassium, eThorium, eTh/K ratio and total magnetic field stacked profiles. Modified from Shives RBK, Charbonneau BE, and Ford KL (2000) The detection of potassic alteration by gamma-ray spectrometry—Recognition of alteration related to mineralization. *Geophysics* 65: 2001–2011.

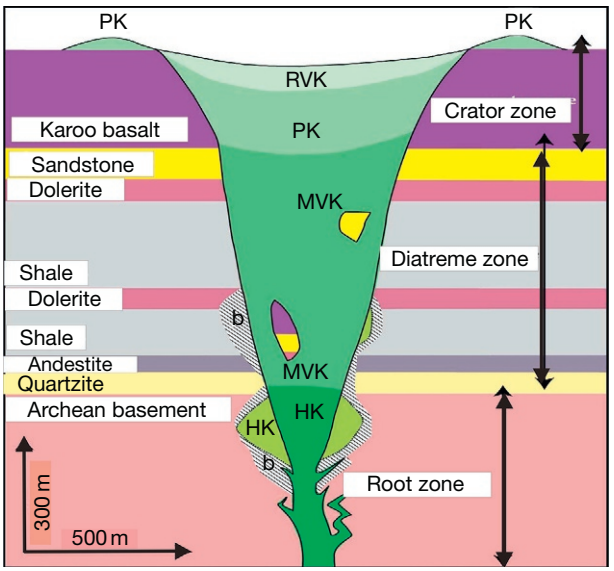


Figure 17 Schematic block diagram illustrating principal elements of kimberlites (root zone, diatreme zone, and crater zone). The geologic formations shown on the left side are from a South African example. Modified from Kjarsgaard BA (2007) Kimberlite pipe models: Significance for exploration. In: Milkereit B (ed.) *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, pp. 667–677.

metamorphic basement rocks. The sandstones are Proterozoic in age, although there are also some Phanerozoic sandstones (Jefferson and Delaney, 2007; Ruzicka, 1995a,b). Faults and fractures within the basement and sandstone units control the location of mineralization and the deposit geometry and orientation. Clay minerals associated with alteration are often associated with the basement and sandstone host rocks (Hoeve and Quirt, 1984). Deposit thickness can be several tens of meters, and lateral extent can be thousands of meters. For example, the Millennium deposit (**Figure 21**) has a minimum strike length of 230 m, a maximum width of 30 m, and a downdip extent of 70 m (Powell et al., 2007). The figure also shows the relationship of the mineralization to the graphitic metasediments and the graphitic marker zone. The mineralized zone is at a depth of 650 m below the surface. Grades of unconformity-related uranium deposits can be >20% U. The two most significant basins for unconformity-related uranium deposits are the Athabasca Basin in Canada (Jefferson and Delaney, 2007) and the McArthur Basin in Australia (Skirrow et al., 2009).

Good overviews of geophysical exploration methods for unconformity-related uranium deposits are provided in Powell et al. (2007) and Ford et al. (2007). **Table 2** in Powell et al. (2007) provides a summary of geophysical indicators for Athabasca uranium deposits. **Figure 21** shows the geologic setting of a typical unconformity-related uranium deposit. The location of the deposit is controlled by faults and lineaments within the basement and overlying sandstone formation. The deposits are intimately related to the graphitic zones that act as a reducing environment for uranium. The graphitic conductors within the basement and overlying sandstone units are excellent conductors. Often, there is clay alteration associated with hydrothermal fluids moving within the basement faults. In addition, the depth to the basement unconformity can be significant (650 m in the case of the Millennium deposit), which has implications on the choice of geophysical methods. For example, a dipole–dipole resistivity array is limited to shallower exploration depths, while a pole–pole resistivity array can map such large conductors at more than 600 m below the ground surface.

Regional magnetic surveys over the Athabasca Basin provide information on basement lithology and structure (Nimick

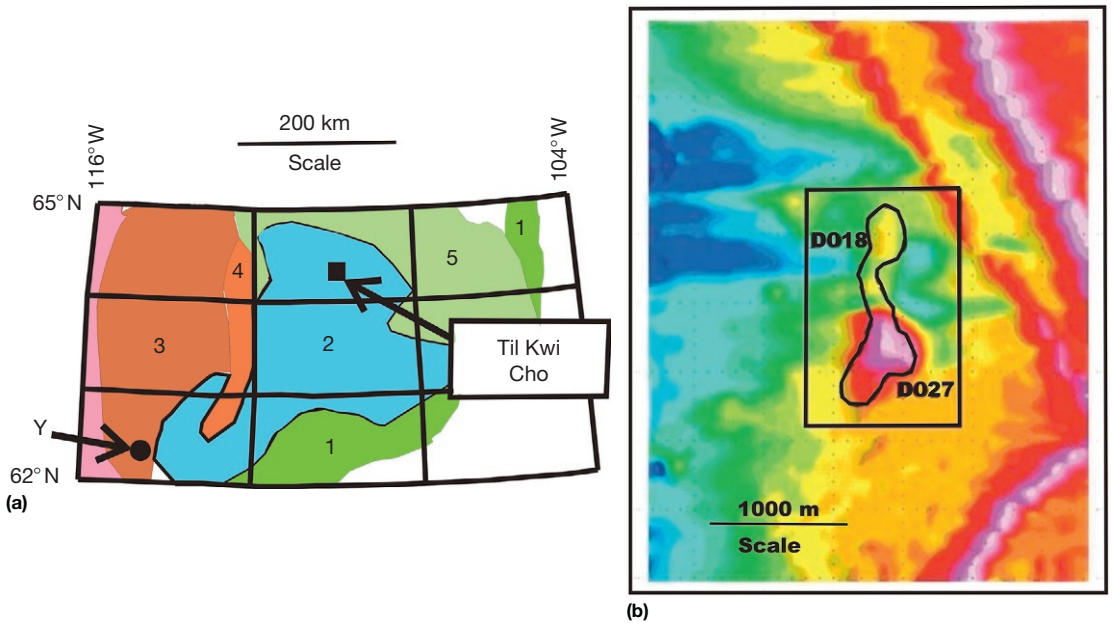


Figure 18 (a) Location map of the Tli Kwi Cho kimberlite complex. (b) Total magnetic intensity map (high values are in red – the D027 anomaly is a +240 nT magnetic high making the range from blue to red ~250 nT) from the 1992 helicopter survey. The black curve outlines the area containing the two kimberlite pipes. Modified from Jansen J and Witherly K (2004) The Tli Kwi Cho Kimberlite Complex, Northwest Territories, Canada: A Geophysical Case Study. *Society of Exploration Geophysicists 74th Annual Meeting, Denver, Colorado*.

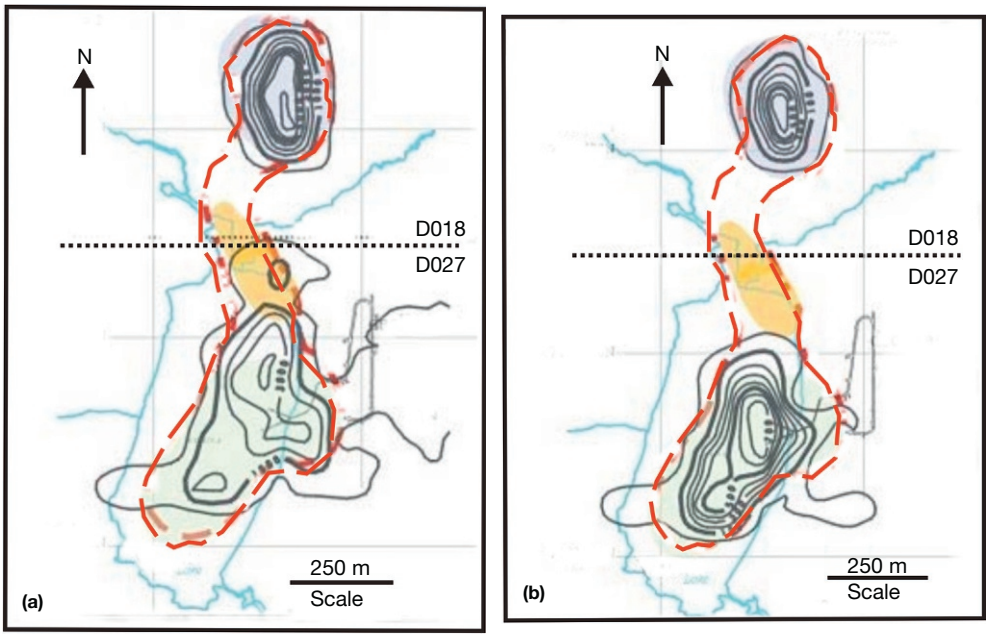


Figure 19 NanoTEM ground EM survey over the Tli Kwi Cho kimberlite complex (modified from Jansen J and Witherly K (2004) The Tli Kwi Cho Kimberlite Complex, Northwest Territories, Canada: A Geophysical Case Study. *Society of Exploration Geophysicists 74th Annual Meeting, Denver, Colorado*). The NanoTEM system has 31 time channels going from ~1 μs to more than 1500 μs. The contour values of (a) and (b) are in mV Am^{-2} with the highest value located in the center of the anomaly. The dashed red line outlines the two pipes. (a) Early time (channel 4 – ~5 μs after turning the transmitter current off: contour interval equals 500 mV Am^{-2} with the central contour (D027) having a value of 3500 mV Am^{-2}). (b) Late time (channel 10 – ~20 μs after the transmitter current is turned off: contour interval equals 100 mV Am^{-2} with the central contour (D027) having a value of 700 mV Am^{-2}).

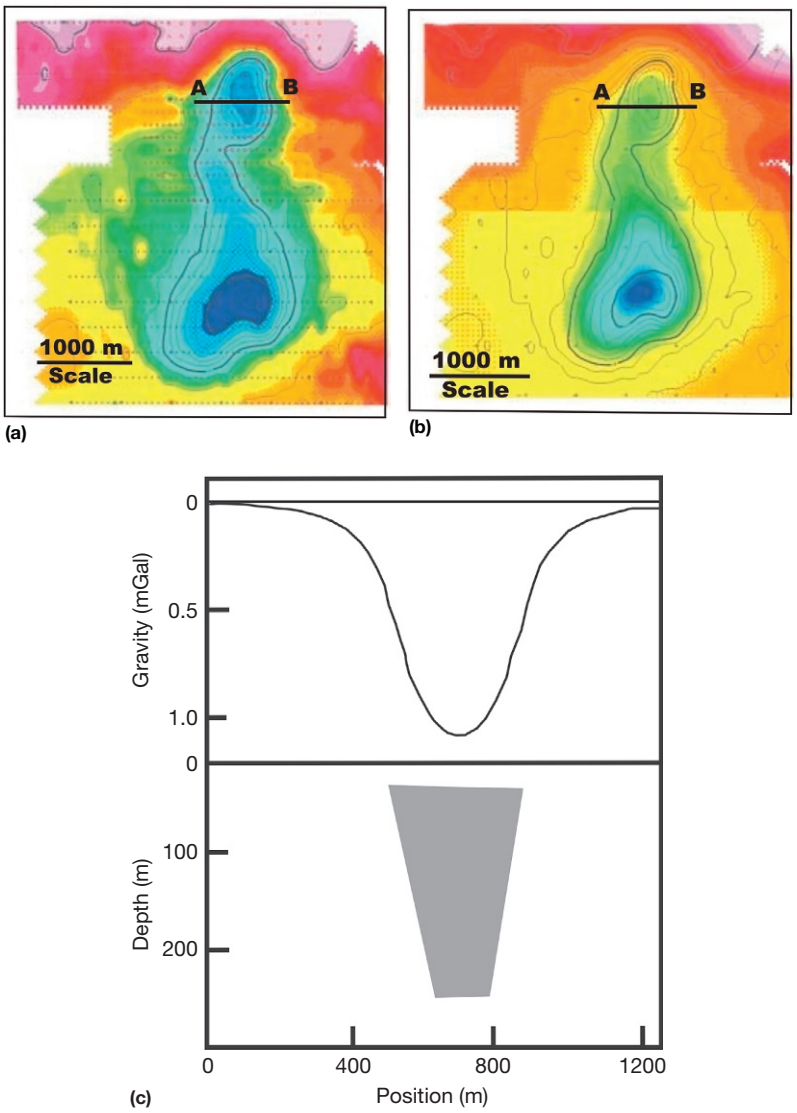


Figure 20 (a) Bouguer gravity data from a ground gravity survey (low values in blue: the Bouguer gravity increases by ~1 mGal) going from blue to red colors. (b) Calculated gravity data using a 2.5-D model. (c) Forward gravity model for the northern kimberlite pipe (D018) along section AB with a density contrast of −300 kg m^{-3}. Modified from Jansen J and Witherly K (2004) The Tli Kwi Cho Kimberlite Complex, Northwest Territories, Canada: A Geophysical Case Study. *Society of Exploration Geophysicists 74th Annual Meeting, Denver, Colorado*.

and Koch, 2008; Powell et al., 2007). **Figure 22** from Nimick and Koch (2008) shows the basement structure and lithology beneath the western portion of the Athabasca Basin. The regional magnetic map in **Figure 23** is near the red star south of the Carswell Formation on the western side of the basin. The magnetic data outline NNW–SSE structures south of the Carswell dome. Such information is important when planning detailed exploration programs.

Magnetic methods provide the regional framework of the basement rocks under the sandstone cover. If there is magnetite associated with basement (or sandstone) faults, it can be mapped using magnetic methods. Gravity methods also provide information on basement structure and are often used in conjunction with magnetic methods. Basement faulting can cause offsets in the basement rocks providing a density contrast that can be mapped using gravity.

The graphitic zones are excellent conductors and have a long strike length. The mineralization is not coincident with the graphitic zones (**Figure 21**) but closely associated with them. Ground EM surveys rely on large loop time-domain EM systems (fixed loop, moving loop, and double loop) and CSAMT and AMT methods (Nimick and Koch, 2008; Smit and Roux, 2009; Tuncer et al., 2006). Deep-penetrating airborne EM methods are required to image through the thick Athabasca sandstone formations (Nimick and Koch, 2008). The EM discussion in the preceding text applies equally well to ground resistivity methods. Dipole–dipole, pole–dipole, and pole–pole resistivity methods are often applied (Legault et al., 2008; Powell et al., 2007). Alteration zones within the sandstone and underlying basement associated with hydrothermal fluids, which may not be good conductors, can be located using resistivity methods. Seismic methods are quite effective

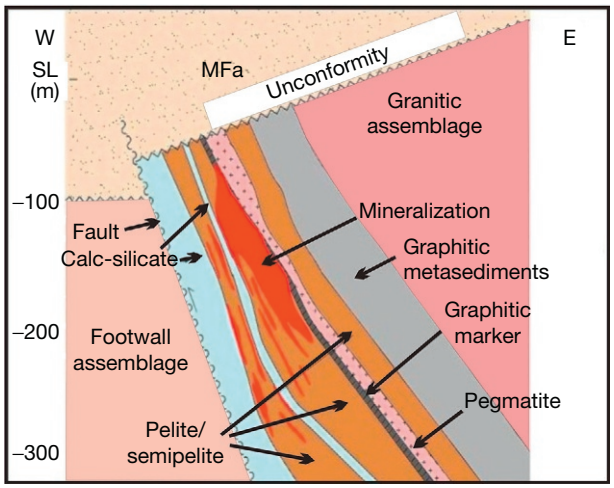


Figure 21 Basement lithologies and location of mineralization of the Millennium deposit, Athabasca Basin. The Athabasca sandstone formation is above the unconformity. Reproduced from Powell B, Wood G, and Bzdel L (2007) Advances in geophysical exploration for uranium deposits in the Athabasca Basin. In: Milkereit B (ed.) *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, pp. 771–790.

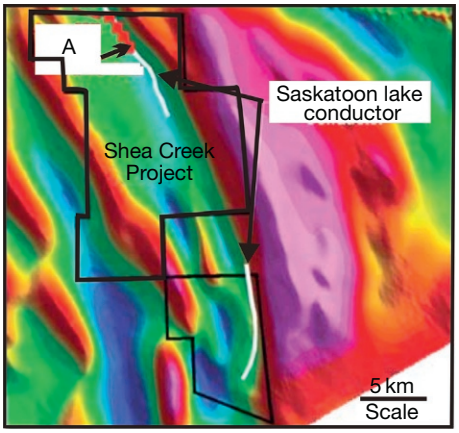


Figure 23 Regional residual airborne magnetic map in the western Athabasca Basin south of the Carswell structure. The purple and red colors represent relatively high magnetic values, and the blue and green colors represent relatively low magnetic values. The Saskatoon Lake conductor lies at the contact between a magnetic high to the east and a magnetic low to the west. The symbol A points to a series of local anomalies associated with the northern portion of the conductor. Modified from Nimick G and Koch R (2008) A progressive geophysical exploration strategy at the Shea Creek uranium deposit. *The Leading Edge*, January, 52–63.

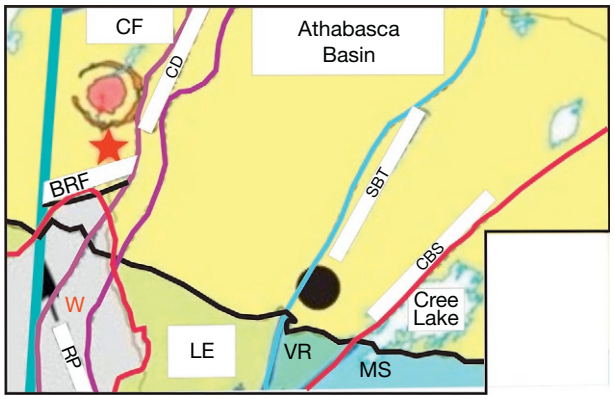


Figure 22 Regional basement structure and geology beneath the western Athabasca Basin. The red star is the Shea Creek zone and the black circle is the Centennial zone. The nearly vertical thick blue line is the Alberta–Saskatchewan border with Lake Athabasca directly north of the outlined area. CF, Carswell Formation; BRF, Beatty River Fault; CD, Clearwater domain (between the purple curves); SBT, Snowbird Tectonic zone; CBS, Cable Bay shear zone; RP, Rae Province; LE, Lloyd (east) domain; VR, Virgin River domain; MS, Mukjatik (south) domain; W, Western Canada Sedimentary Basin. Modified from Nimick G and Koch R (2008) A progressive geophysical exploration strategy at the Shea Creek uranium deposit. *The Leading Edge*, January, 52–63.

in the Athabasca Basin as well (Haijnal et al., 2007; Juhojuntti et al., 2012).

Figures 24 and 25 from Powell et al. (2007) provide an interesting example of using ground EM and resistivity methods to locate potential uranium targets within the Athabasca Basin at ~800 m depth below ground surface. The Centennial zone is located near the southern edge of the Athabasca Basin within the Virgin River domain (black dot in Figure 22). It is bounded on the west side by the Lloyd domain with the contact marked by the Virgin River shear zone (Hoffman, 1990). The Dufferin Lake fault (Figure 24(a)), a reactivation of the Virgin River shear zone, has ~200 m of offset at the unconformity (the northwest side being higher).

Tandem moving loop time-domain EM profiles were carried out using large transmitter loops (1000 × 1000 m), with transmitter–receiver separations of 1000 and 1800 m (line 10+00 in Figure 24(a)). Figure 24(b) shows the late time-normalized voltage data for the late channels (17–20) along line 10+00N. Two anomalies associated with deep conductors can be seen at ~19 900 and 21 800 easting. The one at 19 900 lines up with the C1 conductor.

A pole–pole resistivity survey was carried out along three lines parallel to the C1 conductor, that is, perpendicular to line 10+00N (a = 200 m, n = 1–10) with a line spacing of 500 m. A 3-D inversion of the data along the three lines was then carried out using the Res3dinv inversion code (Loke and Barker, 1996a,b). Figure 25(a) is the section closest to the C1 conductor showing two zones of lower resistivity (labeled breaches – thought to be related to structure and alteration of the sandstone due to hydrothermal fluids) within the highly resistive (> 10 000 Ω-m) sandstone.

Two holes were drilled based on the EM and resistivity data. The discovery hole VR-18 is located on the north side of the southern breach (Figure 25(a)). This hole intersected perched uranium in the sandstone from 710.5 to 711.7 m and 752.0 to 762.8 m and unconformity-hosted uranium from 789.1 to 795.5 m, with average grades of 0.27%, 1.0%, and 5.83% U_3O_8, respectively (Powell et al., 2007).

Figure 25(b) displays the borehole resistivity data for hole VR-18. There is a high-resistivity band (2000 Ω-m) near the surface, and in several bands, depth is between 230 and 570 m with moderate resistivity values (1500 Ω-m) elsewhere, except

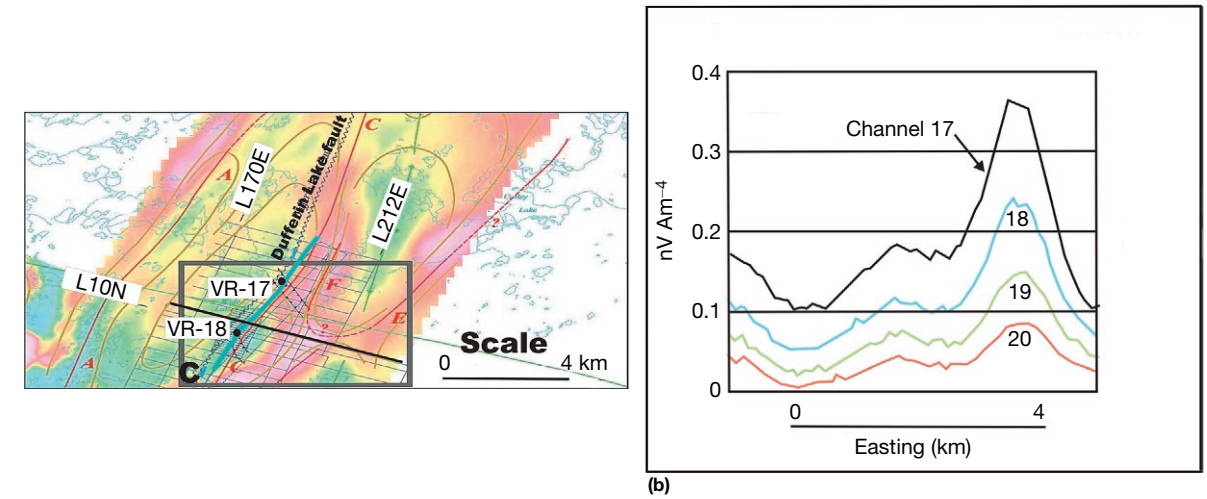


Figure 24 (a) Centennial zone aeromagnetic vertical gradient map with interpreted basement fold saddle (box area). The red colors are positive gradient values and the blue are negative gradient values. Moving loop data along line L10+00N (black line), and discovery hole VR-18. The C1 conductor (labeled C and shown as a blue line) parallels the Dufferin fault on the down-thrown side. (b) Late time voltage channels from the moving loop EM survey with the transmitter–receiver separation equal to 1800 m. The survey was conducted along the black line (Line 10) shown in (a). Modified from Powell B, Wood G, and Bzdel L (2007) Advances in geophysical exploration for uranium deposits in the Athabasca Basin. In: Milkereit B (ed.) *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, pp. 771–790.

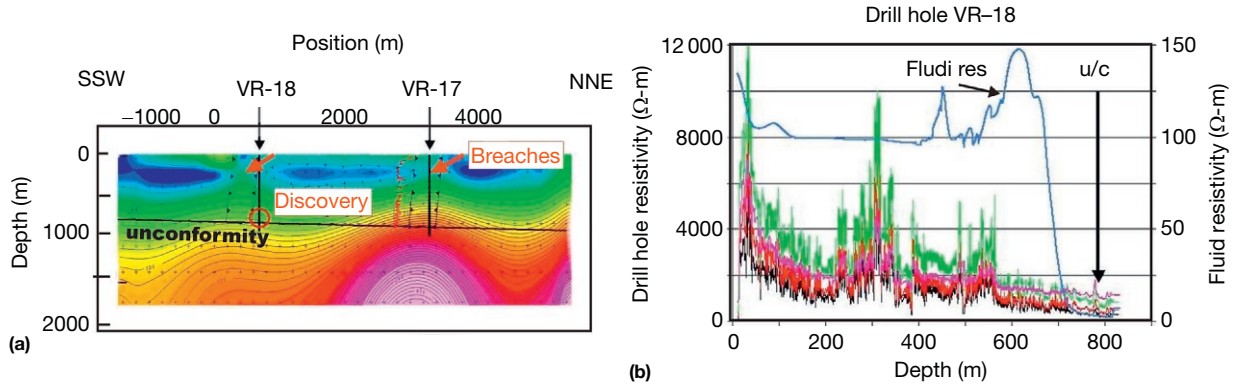


Figure 25 (a) Longitudinal conductivity section along conductor C (blue line parallel to the conductor in Figure 24a – labeled with the black C). The resistivity range for the pole–pole-survey is 10 000 (blue)–100 Ω-m (pink). The sloping black line is the unconformity contact. VR-17 and VR-18 are the locations of the two boreholes. (b) Normal resistivity log versus depth (the colors represent different separations between the borehole transmitter and receiver electrodes) and water resistivity (in blue).The unconformity is located at u/c. Modified from Powell B, Wood G, Bzdel L (2007) Advances in geophysical exploration for uranium deposits in the Athabasca Basin. In: Milkereit B (ed.) *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, pp. 771–790.

near the bottom where the sandstone resistivity is < 1000 Ω-m. These results are roughly comparable to the pole–pole resistivity values. The fluid resistivity in borehole VR-18 (blue curve in Figure 25(b)) remains high (~ 100 Ω-m) indicating freshwater to a depth of ~ 680 m from surface. Near the bottom of the hole, the resistivity drops to very low values indicating brine resting on the sandstone–basement unconformity. This is a possible explanation of the C1 conductor from the EM survey.

11.15.4.2.5 Laterite (nickel) deposits

Laterite is a surficial formation that forms in hot, wet tropical areas of the world (Aleva, 1994; Golightly, 1981). Most rocks can be deeply weathered from exposure to high rainfall and elevated temperatures. The rainwater penetrates the rock exposed at the surface and leaches out easily soluble elements such as sodium, potassium, calcium, magnesium, and silicon. This leads to a residual concentration of more insoluble elements such as iron and aluminum. Laterites consist mainly of the minerals kaolinite, goethite, hematite, and gibbsite, which form in the course of weathering. Many laterites also contain quartz since it can be a relatively stable mineral of the parent rock.

Weathering of ultrabasic igneous rock (serpentinite, peridotite, and dunite containing low concentrations of nickel – 0.2–0.3%) can lead to concentration of nickel in nickel-rich oxides and silicates (Golightly, 1981). These surficial deposits are typically a few meters thick but can be thicker depending on surface topography (Figure 26).

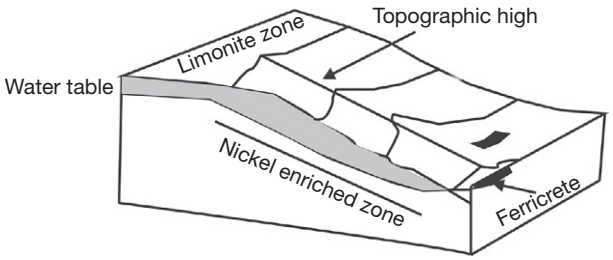


Figure 26 Schematic of a potential lateritic nickel environment.

The resistivity difference between the weathered zone and the underlying bedrock may be large enough that resistivity methods can map the contact between. Dielectric constant differences between the different weathered zones and bedrock allow GPR to map these zones and the bedrock contact (Francke, 2012). Multichannel analysis of surface wave techniques (Miller et al., 1999; Park et al., 1999) has been used to map the thickness of the weathered zone.

Figure 27 is an example of a GPR profile over a nickel laterite zone from Francke (2012). Figure 27(a) is a schematic showing the different zones within the deposit. Figure 27(b) is the processed 30 MHz GPR profile with topographic relief. The boreholes overlain on the profile outline the limonite and saprolite zones (note colors on boreholes). Note the absence of reflectors within the limonite zone and the bedrock beneath the rocky saprolite zone.

11.15.4.2.6 Placer deposits

Placer deposits are the result of the concentration of heavy minerals that have been released from their parent rock. These heavy minerals are moved by water into streams where they are concentrated in gravel beds (Morison, 1989; Sutherland, 1991). The minerals must be heavier than quartz and must resist weathering processes. The main minerals and metals associated with placer deposits are gold, cassiterite (tin), platinum, diamonds, magnetite, and some gemstones. The two most important types of placers are stream (alluvial) and beach placers. The geophysical methods used for aggregate (sand and gravel) mapping are identical to the geophysical methods used for mapping placer deposits. The concentrations of the metals and/or minerals in placer deposits are too low for their direct detection using geophysical methods. However, properties of the unconsolidated sediments associated with placers can be determined using geophysical methods.

The velocity of the unconsolidated sediments containing the placer deposit is lower than the underlying bedrock velocity making refraction seismic an effective method of determining the thickness of the sediments. The resistivity of the unconsolidated sediments containing the placer deposit is generally lower than the underlying bedrock; consequently, EM methods can be effective for determining the thickness of these sediments. The same resistivity contrast discussed in the preceding text for EM methods applies to resistivity methods. In addition to mapping, the total sediment thickness in certain cases resistivity methods can also determine the gravel layers from glacial tills and sand layers. The dielectric constant of the unconsolidated sediments is often different than the underlying bedrock. If the unconsolidated sediments are not conductive, then GPR methods can be used to map the sediment thickness and to map individual layers within the sediment package.

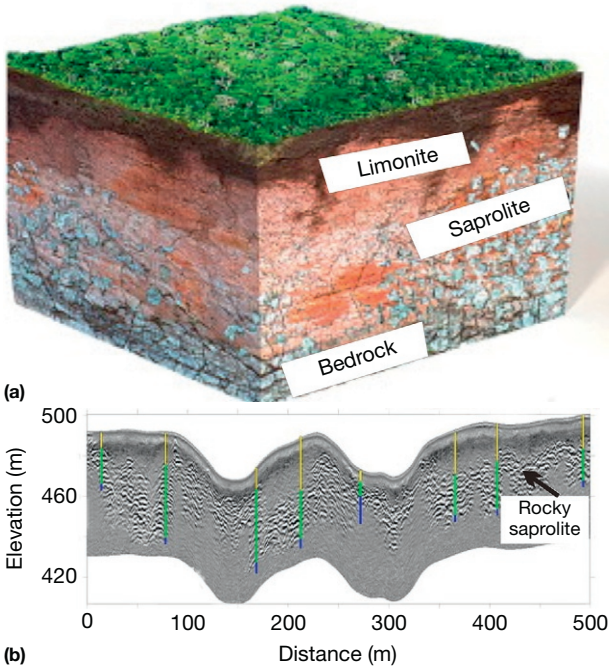


Figure 27 (a) Model of a nickel laterite profile. (b) 30 MHz radar profile showing the occurrence of rocky saprolite (green on boreholes) within the profile (the velocity used to convert time to depth is 0.067 m ns^{-1}). The absence of reflectors above this zone is characteristic of early limonite and saprolite. Beneath the rocky saprolite zone is bedrock, which exhibits no radar reflectors. Borehole color codes: yellow, limonite; green, saprolite; blue, bedrock. 2:1 vertical exaggeration. Modified from Francke J (2012) A review of selected ground penetrating radar applications to mineral resource evaluations. *Journal of Applied Geophysics* 81: 29–37, with permission from Elsevier Scientific Publishing.

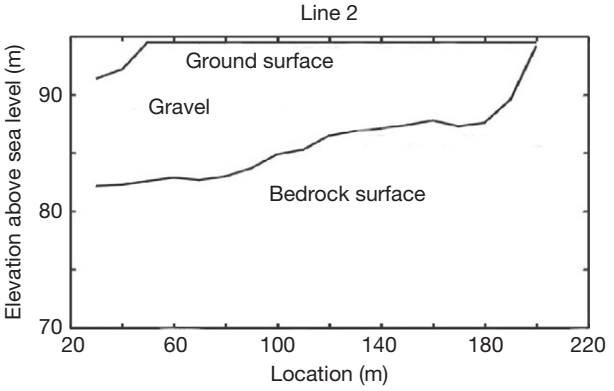


Figure 28 The bedrock topography and gravel thickness were computed from a ray-tracing program using the TX plots from the refraction seismic survey.

Figure 28 is an example of mapping gravel thickness using refraction seismic. The data were collected at a gravel deposit on Vancouver Island, Canada. The operator was interested in

determining if there was enough gravel left to warrant further mining. A 12-channel Geometrics instrument was used with a geophone spacing of 10 m. The hammer source was located at 10 and 50 m off both ends of the 110 m spread. The first breaks for each 110 m spread were plotted as a time versus position (T–X) plot. These data were then used in a ray-tracing program to compute the gravel thickness as a function of position. The gravel velocities for the three profiles collected were 697, 881, and 867 m s^{-1} for profiles 1, 2, and 3, respectively, which are consistent for dry gravel. The bedrock velocities for the three profiles were 5904, 4332, and 6716 m s^{-1} for profiles 1, 2, and 3, respectively, which are consistent for volcanic rocks. Figure 28 shows the interpreted profile for line 2. The bedrock outcrops at approximately station 200, which is in agreement with the profile. The gravel was ~10 m thick, which the operator decided to be not enough to warrant further mining.

Figure 29 is an example of mapping gravel thickness using time-domain EM methods. The object in this case was to map the gravel thickness of a placer gold deposit in southern Chile, South America. The data were collected using the Geonics EM-47 with a 40 m square transmitter loop. The system was used in the central sounding mode with the receiver at the center of the transmitter loop, and a sounding was collected every 80 m along the lines. The Interpex 1-D inversion software (Interpex, 2006) was used to edit and interpret the data. The left side of Figure 29(a) shows the measured data (symbols) and the calculated data from the best fit two-layer model (black solid line). The calculated data were computed using the solid black two-layer model shown on the right side of Figure 29(a). The colored lines shown in Figure 29(a) represent other two-layer models that have the same error of fit as the best fit model. Consequently, the depth error is between 5 and 8 m with a best fit depth of ~35 m. Figure 29(b) shows two profile plots of overburden thickness. The plotted thicknesses are the best fit value from each sounding along the profiles. The two test holes (holes 4 and 6) dug to verify depth to bedrock show how well the EM method was able to map gravel thickness. Figure 29(c) is a plot of overburden thickness covering the entire survey area. The blue colors represent thin overburden thickness, and the purple colors represent thick overburden thickness. Note how a ridge (blue colors) separates two thick gravel sections on each side.

11.15.4.2.7 Other deposits

SedEx Pb–Zn–Ag deposits are hosted in marine sedimentary rocks within intracratonic or epicratonic rift basins. Intracratonic basins lie within the continental crust, for example, the Williston Basin in North America. Epicratonic basins lie on the continental crust but are partially open to an ocean basin, for example, the Niger Basin of West Africa. Deposits occur in carbonaceous shales in basin carbonate, shale, or siltstone facies that were deposited on thick sequences of rift-fill conglomerates, red beds, sandstones or siltstones, and mafic or felsic volcanic rocks. Ore bodies are generally tabular or stratiform and are localized in smaller fault-controlled subbasins near the margins of major depocenters. Laminated or bedded sulfide ores that are hosted in carbonaceous and pyritic, fine-grained shales and siltstones define this deposit type (Emsbo, 2009; Lydon, 1995). The main economic minerals are sphalerite and galena; however, chalcopyrite may be a minor component. Pyrite is the most abundant sulfide, though pyrrhotite may also be present.

With the exception of the Sudbury Basin, which was formed by a meteorite impact, magmatic Ni–Cu–PGE deposits are found in mafic and ultramafic igneous intrusions and volcanic flows. Mafic rocks include basalt, dolerite, and gabbro, and ultramafic rocks include dunites, peridotites, pyroxenites, komatiites, and anorthosites (Eckstrand and Hulbert, 2007; Naldrett, 2004). The metals in these deposits are contained in sulfide minerals of Ni, Cu, and PGEs. Ni grades are typically between 0.7% and 3%, and Cu grades are between 0.2% and 2%. Ore tonnages of individual deposits range from a few hundred thousand to a few tens of millions (Eckstrand and Hulbert, 2007). Two giant Ni–Cu districts stand out above all the rest in the world: Sudbury, Ontario, and Norilsk-Talnakh, Russia, with ore tonnages of 1645 and 1903 million tons, respectively (Eckstrand and Hulbert, 2007).

SedEx and magmatic Cu–Ni PGE deposits have many characteristics similar to VMS deposits. Both these deposit types are amenable to EM (resistivity) methods due to the high concentration of sulfide minerals. In fact, magmatic Cu–Ni deposits are extremely conductive; for example, the Voisey's Bay deposit in Labrador, Canada, has a conductance (conductivity multiplied by the width of the deposit) of several thousand siemens (Balch, 1999). These deposits often have a strong magnetic response due to the presence of pyrrhotite and can have a positive density contrast resulting from the massive sulfide minerals (such that gravity methods can be effective). Radiometric methods can map sericitic alteration in the form of elevated potassium values and depleted thorium/potassium ratios associated with SedEx deposits relative to those of the host rocks (Shives et al., 2000).

Sedimentary deposits are deposits that are located within sedimentary rocks. The two main types of these deposits are carbonate-hosted Pb–Zn deposits (Mississippi Valley-type deposits) and sediment-hosted stratiform Cu deposits. The Pb–Zn deposits are stratabound limestone or dolomite deposits in carbonate platform sequences (Sangster, 1995). They generally occur in clusters and tend to occupy voids in carbonate breccias. Sediment-hosted stratiform Cu deposits (red bed Cu deposits) are stratiform deposits hosted in anoxic marine sediments usually above red oxidized continental clastic sediments (Kirkham, 1984, 1989).

Gravity (high density of sphalerite and galena), resistivity (conductivity of galena and associated minerals such as pyrite and pyrrhotite), and IP (more disseminated Pb–Zn sulfide mineralization) are used to locate carbonate Pb–Zn deposits containing more massive sulfide mineralization (Lajoie and Klein, 1979; Seigel et al., 1968). Resistivity (high conductivity of chalcopyrite) is commonly used to locate stratiform Cu deposits. Seismic methods have been used to locate both carbonate-hosted Pb–Zn and stratiform Cu deposits (Hearst et al., 1994; Lajoie and Klein, 1979). If there is sufficient pyrrhotite within these deposits, magnetic methods may be a useful exploration method.

Hydrothermal gold deposits are formed in environments where hot geothermal water carrying dissolved gold is focused into a confined space, where there is some mechanism for precipitation of the gold (Figure 30). Examples are fault and shear zone deposits, intrusion-related deposits, and epigenetic

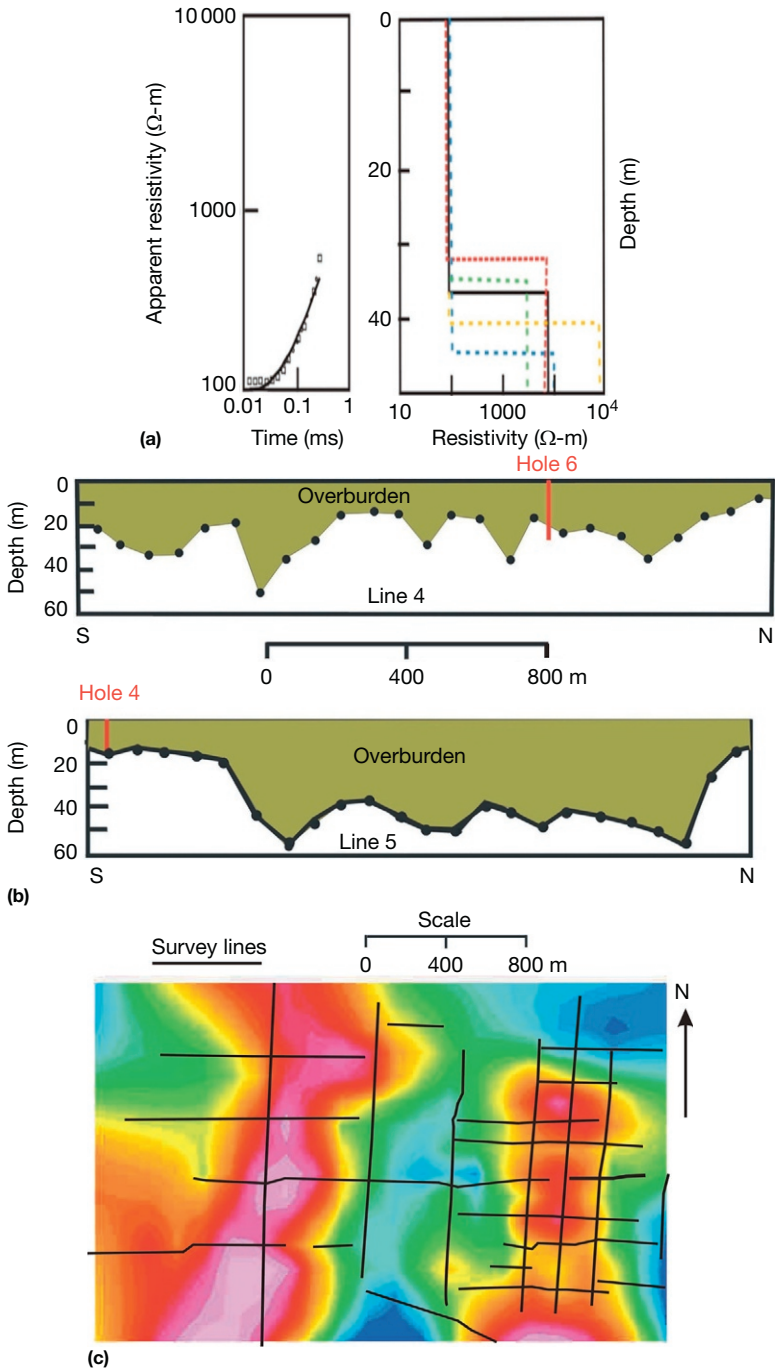


Figure 29 (a) One-layer EM-47 inversion for line 5, station 800. The data are on the left side and the best fit model (solid black line) and range of fits (colored lines) with the same error of fit are on the right side. (b) Two profiles computed using the best fit depths from inversion with two test holes dug to verify the interpretation. (c) Map of the overburden thickness showing a ridge (blue colors) in the middle, which separates two thick gravel packages (purple and red colors) on each side. The solid black lines show where the EM soundings were taken.

(high and low sulfidation) deposits (Lydon et al., 2004; Williams, 1997). These hot fluids often cause alteration of the country rock within the vicinity of the deposit.

Gold, like diamonds, does not occur in large enough concentration to be directly detected using geophysical methods. However, a variety of geophysical techniques are applicable to gold exploration, with the specific technique(s) depending on lithologic, mineralogical, and alteration characteristics of the deposit type (Ford et al., 2007; Irvine and Smith, 1990; Johnson and Fujita, 1985; Williams, 1997). Magnetic data can delineate different lithologies, regional faulting patterns, and shear zones, particularly for shear- and fault-zone (Au-quartz vein) deposits. EM methods are sometimes used to map faults, veins, contacts, and alteration. Resistivity/IP and radiometric

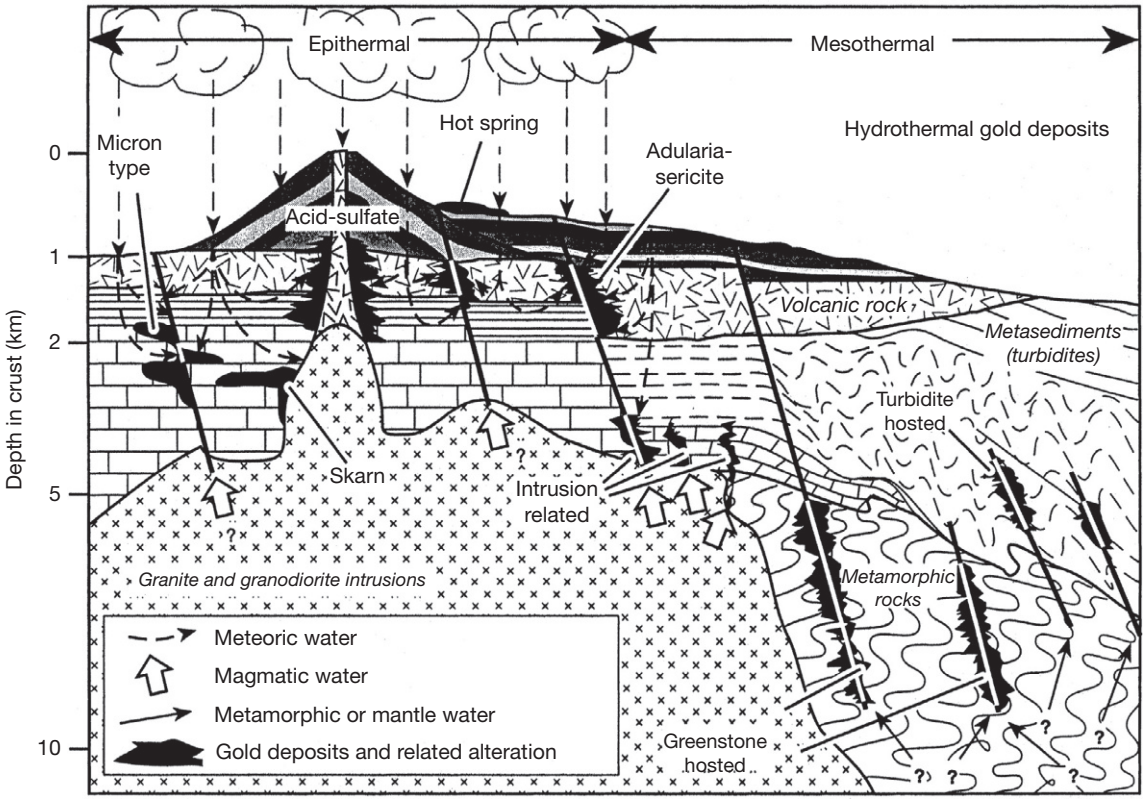


Figure 30 Depositional environments for hydrothermal gold deposits. Modified from Williams PK (1997) Towards a multidisciplinary integrated exploration process for gold discovery. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral Exploration*, pp. 1015–1028.

methods can be used to map larger quartz veins (resistivity highs) and potassium (high) alteration.

There are several geophysical techniques that can be applied to delineate structures and alteration for epithermal deposits (Ford et al., 2007; Hoover et al., 1992; Kawasaki et al., 1986; Rona, 1978). Regional gravity can map gravity highs associated with felsic intrusions. Magnetic lows are often associated with alteration products. Regional potassium highs can be associated with felsic volcanism, and as noted earlier, local potassium highs with corresponding low Th/K ratio are associated with potassic alteration. Resistivity lows may be associated with sulfides, argillic alteration, or perhaps increased porosity. Resistivity highs can help recognize regions associated with silicification and/or intrusions.

An interesting deposit type is the carbonate replacement (CRD) style associated with Ag, Pb, and Zn mineralization (Robertson and Megraw, 2009). CRDs form a spectrum of styles ranging from stock contact skarns to dike and sill contact skarns and massive sulfides to massive sulfide chimneys and mantos, coherent massive sulfide sheets (Einaudi et al., 1981; Meinert, 1982). Geophysical methods commonly employed for this type of deposit are EM and resistivity (massive sulfides), gravity (massive sulfides), and magnetic methods (pyrrhotite, magnetic high, and alteration, magnetic low). IP methods, combined with resistivity surveys, are useful for mapping Pb–Zn mineralization. Both airborne and ground geophysical methods have been applied for this type of deposit.

11.15.5 Nonmetallic Deposits

Nonmetallic deposits form a class of deposits that are associated with nonmetallic minerals and/or deposits without metal being the product of interest. They are an important class of mineral deposits. **Table 8** is a partial list of nonmetallic commodities of economic interest. The list includes coal, potash (K_2O and KCl), salts (NaCl, KCl, and KNO_3), barite ($BaSO_4$), cement (limestone – $CaCO_3$), gypsum ($CaSO_4.2H_2O$), phosphate (apatite), clays, sand and gravel, and graphite. Coal, potash, salt, barite, limestone, gypsum, and most deposits of phosphate are associated with sedimentary rocks. Clays and sand and gravel are associated with unconsolidated surficial sediments. Graphite is associated with metamorphic and igneous rocks but is not a metal.

Potash, salts, gypsum, and barite are evaporites, that is, formed by evaporation of seawater in a restricted basin (Melvin, 1991). Geophysical techniques used to explore and delineate these deposits and the other sedimentary deposits listed in **Table 8** are generally the same as those used for oil and gas exploration, for example, 2-D and 3-D seismic, gravity, and electrical resistivity tomography (ERT – used for mapping bitumen grade of oil sands) and borehole logging (Gendzwill and Brehm, 1993; Nicol and Lawton, 2012; Pesowski and Larson, 2000; Tixier and Alger, 1970). GPR and gravity methods are also used during mining operations of evaporite deposits (Annan et al., 1988; Kovin, 2010; Pringle et al., 2012).

Table 8 Nonmetallic deposits

Commodity	*Mineral or chemical formulae*	*Comments*
	Sedimentary deposits	
Coal	Mainly carbon with impurities	Used for fuel and coke
Potash	K_2O, KCl	Used for fertilizer
Salts	NaCl, KNO_3	Industrial uses, food preservatives, fertilizer, deicing, table salt
Barite	$BaSO_4$	Used as drilling additive
Cement	Limestone ($CaCO_3$)	Used for concrete and plaster
Gypsum	$CaSO_4.2H_2O$	Used for wall board
Phosphate	Apatite ($Ca_{10}(PO_4)_6(OH,F,Cl)_2$)	Apatite is the main source of phosphate. Used mostly for fertilizer
	Surficial deposits	
Various clays	Kaolinite [$Al_2Si_2O_5(OH)_4$], montmorillonite [$(Na,Ca)_{0.33}(Al,Mg)_2(Si_4O_{10})(OH)_2nH_2O$], illite [$(K,H_3O)(Al,Mg,Fe)_2(Si,Al)_4O_{10}\{(OH)_2,(H_2O)\}$]	Used for pottery, building material
Sand and gravel		Used in concrete and for construction
	Other deposits	
Graphite	Pure carbon	Used for, lubricants, batteries, steel making

Fracture detection and delineation are very important for potash and other evaporite-mining operations. Fracturing increases the probability of failure of the supporting pillars and the formations above the evaporite zone that prevents flooding of the mine with water. The GPR method is quite effective for mapping fractures within evaporites such as potash and halite because these formations are resistive, thus allowing significant penetration of the GPR signal. **Figure 31** is an example from Kovin (2010) using a 400 MHz GPR system within a Russian potash mine. **Figure 31(a)** and **31(b)** is the 2-D raw and processed data showing the direct wave, mine wall reflection, and two fractures. **Figure 31(c)** is a 3-D volume showing the orientation of the fractures.

Facture mapping has also been carried out from the surface using 3-D seismic surveying. Nicol and Lawton (2012) showed an example of 3-D three-component seismic surveys carried out over a producing potash mine (within the Prairie Evaporite Formation) in Saskatchewan. Room and pillar mining was employed that led to stresses and fracturing within an overlying carbonate (Dawson Bay Formation). In this case, near-vertical fractures in the relatively flat carbonate strata are expected to reduce seismic velocity (both *S*-wave and *P*-wave) perpendicular to the orientation of fracturing. **Figure 32** is a seismic section with the interval V_p/V_s ratio overlain in color. The V_p/V_s ratios between the Birdbear Formation and the Dawson Bay Formation are representative of nonfractured carbonates with V_p/V_s ratios between 1.76 and 1.97. The blue strip of V_p/V_s ratios between the top of the Dawson Bay Formation and the top of the Prairie Evaporite Formation is between 1.99

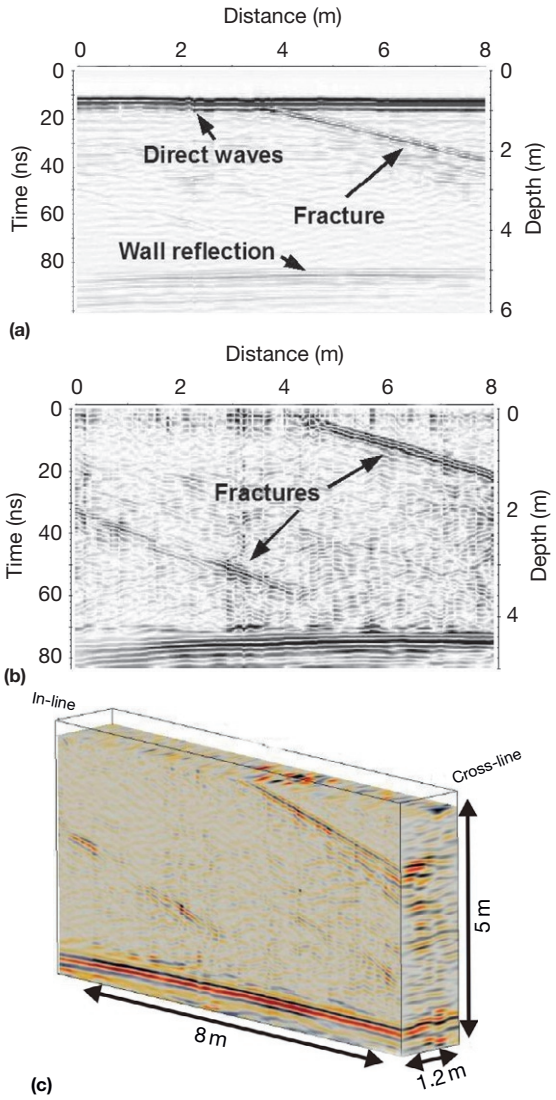


Figure 31 (a) Raw record of 2D GPR data in a potash mine (the conversion of time to depth was carried out using a velocity of 0.12 m ns^{-1}). (b) Processed data; note that after processing, a prominent reflection from a second fracture is clearly identified. (c) Three-dimensional data volume calculated using the 2D GPR profiling data. The reflections from visible and suspected fractures are clearly identified. Modified from Kovin ON (2010) Fracture imaging in the upper Kama potash mine using 3D GPR data. *Expanded Abstract, Symposium on the Application of Geophysics to Environmental and Engineering Problems, Keystone Colorado*.

and 2.31, indicating an increase in V_p/V_s ratio within the Dawson Bay Formation, which corresponds to fracturing in this formation. Once into the Prairie Evaporite Formation, V_p/V_s ratios drop to a range of 1.40–1.97 as expected.

Gravity surveying is employed for the exploration and development of halite (NaCl) and potash deposits since they have relatively low density (approximate densities of halite and potash are 1900 and 1281 kg m^{-3}, respectively). An example of using gravity methods for long-term monitoring of an abandoned salt mine in England is given in a paper by Pringle et al. (2012). **Figure 33(a)** is a plot of repeat gravity surveys over

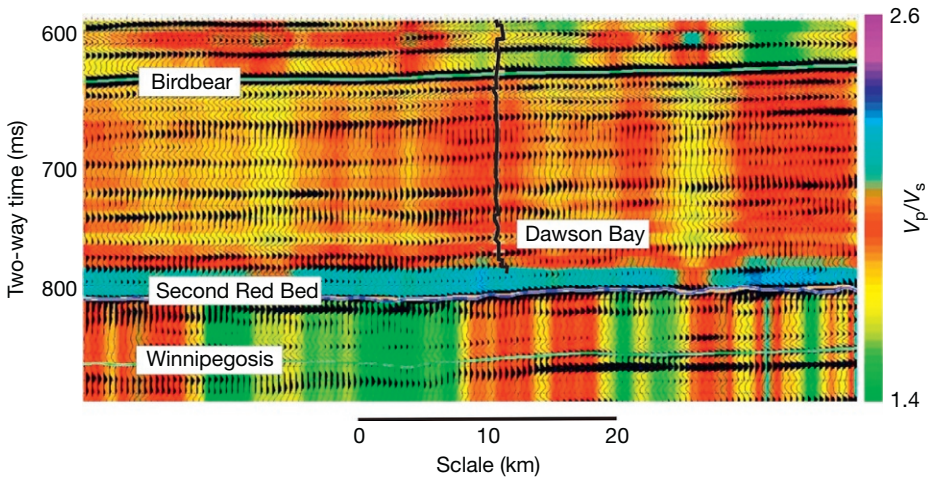


Figure 32 Seismic cross section with interval V_p/V_s overlain in color from a potash area in the Saskatchewan portion of the Williston Basin Saskatchewan. Well control is shown as the black nearly vertical line near the center of the section. Modified from Nicol A Lawton DC (2012) Multi-component investigation of fracturing in a potash mining region in Saskatchewan, Canada. *Expanded Abstracts, Society of Exploration Geophysics Annual Meeting, Las Vegas, Nevada*.

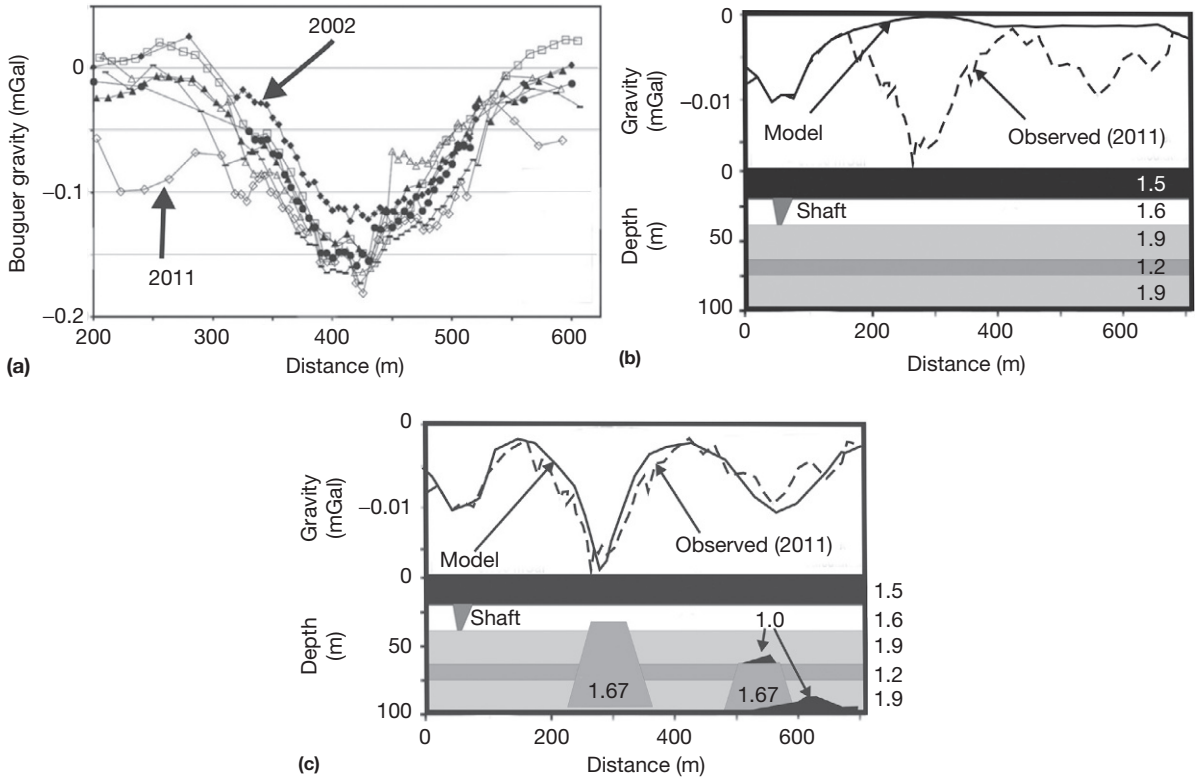


Figure 33 (a) The change in Bouguer gravity (mGal) over an abandoned mine during the period from 2002 to 2011. (b) Gravity model compared with the observed 2011 data with no collapsed material other than the collapsed mine shaft at the left end (densities are given in g cm^{-3}) and (c) gravity model compared with the observed 2011 data with collapsed material (density of 1.67 g cm^{-3} – 1670 kg m^{-3}). The improvement in fit between the measured Bouguer gravity and computed Bouguer gravity is evident. Modified from Pringle JK, Styles P, Howell CP, Branston MW, Furner R, and Toon SM (2012) Long-term time-lapse microgravity and geotechnical monitoring of relict salt mines, Marston, Cheshire, UK. *Geophysics* 77: B287–B294.

the abandoned salt mine from 2002 to 2011. Note how the minimum Bouguer gravity value decreased from approximately −0.12 to approximately −0.17 mGal during that time indicating an increase in collapsed material. Figure 33(b) models the abandoned mine with a collapsed mine shaft entrance on the left but no collapsed regions elsewhere. The modeled data do not agree with the measured data, even when using density values from logs and information on elevations within the mine. The model in Figure 33(c) was altered from Figure 33(b) to include two near-surface voids and associated collapse material areas centered at ~290 and 550 m, respectively. Modeled brine-filled voids with a

density of 1000 kg m^{-3} were located at ~25 and 60 m below ground level, respectively. Underlying modeled collapsed material extended down to mine workings (~90 m below ground level) and had a relatively low density of 1670 kg m^{-3} since the material usually expands after collapse.

The geophysical methods described in the preceding text are also used to explore for sedimentary phosphate deposits. Nonsedimentary phosphate deposits are also associated with carbonatites, carbonate igneous intrusions (Bell, 1989; Kresten, 1983). The geophysical techniques used to explore for them are similar to those used for kimberlite exploration. Gravity, ERT, and seismic methods are the geophysical methods employed for coal exploration. Since graphite is an excellent conductor, EM and resistivity techniques are the geophysical methods of choice.

11.15.6 Recent Geophysical Developments

Recent improvements in airborne gravity and the development of airborne gravity gradiometry are now being applied for regional mapping and for direct detection of mineral targets. During the last two decades, 2-D and 3-D reflection seismic methods developed for petroleum exploration have been modified for hard-rock mineral exploration. The application of these two methods to mineral exploration will be presented in the succeeding text.

11.15.6.1 Airborne Gravity and Gravity Gradiometry

Although airborne gravity has been in use for many years, recent developments in hardware, processing, and interpretation have significantly improved resolution (Argyle et al., 2000; Forsberg and Olsen, 2010; Sander et al., 2005; Verdun et al., 2002). The following example illustrates the significant improvements in airborne gravity. Geoscience BC's QUEST Project is located in central British Columbia, Canada. The object of the project was to explore beneath thick glacial cover. The geophysical portion of the project consisted of flying EM, magnetic, and gravity surveys. The airborne gravity survey was flown using the Sander AIRgrav system with a line spacing of 2 km. The gravity data provided information on depth of glacial cover, bedrock lithology, structure, and alteration products based on rock density. The Geological Survey of Canada (GSC) carried out a ground gravity survey at an approximate station spacing of 10 km (**Figure 34(a)**); the AIRgrav survey therefore had ~25 times the coverage. The AIRgrav Bouguer gravity data (**Figure 34(b)**) were processed using a density of 2670 kg m^{-3} and a filter length of 4 km. As expected, there is a significant increase in detail with the AIRgrav data.

The first airborne gravity gradiometer system was developed by BHP Billiton in the mid 1990s and early 2000s using technology developed for the US Navy by Bell Geospace Inc. – now Lockheed – Martin (Dransfield and Lee, 2004; Hammond, 2003; Lee, 2001; van Leeuwen, 2000). Since then, several other commercial systems have come on the market. Traditional gravity systems measure the vertical component of the Earth's gravitational field (the unit is the Gal = 1 cm s^{-2}). A tensor gradiometer system measures the components of the gravity gradient tensor. It is a rank 3 tensor with nine components ($\partial g_x/\partial x = G_{xx}$, $\partial g_x/\partial y = G_{xy}$, $\partial g_x/\partial z = G_{xz}$, ..., $\partial g_z/\partial x = G_{zx}$, $\partial g_z/\partial y = G_{zy}$, and $\partial g_z/\partial z = G_{zz}$ with g_x, g_y, and g_z representing the three components of the gravitational field). However, there are only five independent components because the sum of the diagonal elements turns out to be zero and from symmetry $G_{xy} = G_{yx}$, $G_{xz} = G_{zx}$, and $G_{yz} = G_{zy}$. The gravity gradient has units of gravity acceleration (Gal) per unit distance and is measured in Eötvös (Eo) with 1 Eo = 10^{-7} Gal m^{-1} = 10^{-4}mGal m^{-1} = 10^{-9} s^{-2}.

Most of the focus to date has been on the vertical gradient of the vertical gravity acceleration G_{zz}. **Figure 35** from Mims and Mataragio (2010) is an example showing how independent tensor components respond to a body with a positive density

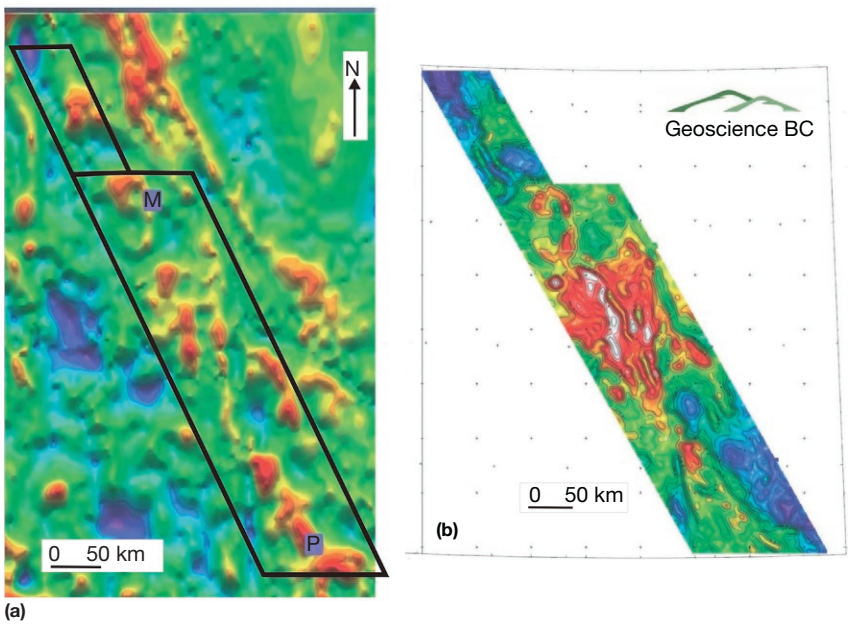


Figure 34 (a) GSC ground gravity at an approximate spacing of 10 km. The QUEST Project area is outlined with black lines. (b) The AIRgrav Bouguer gravity computed using a density of 2670 kg m^{-3} and a filter length of 4 km. Modified from Geoscience BC's QUEST Project.

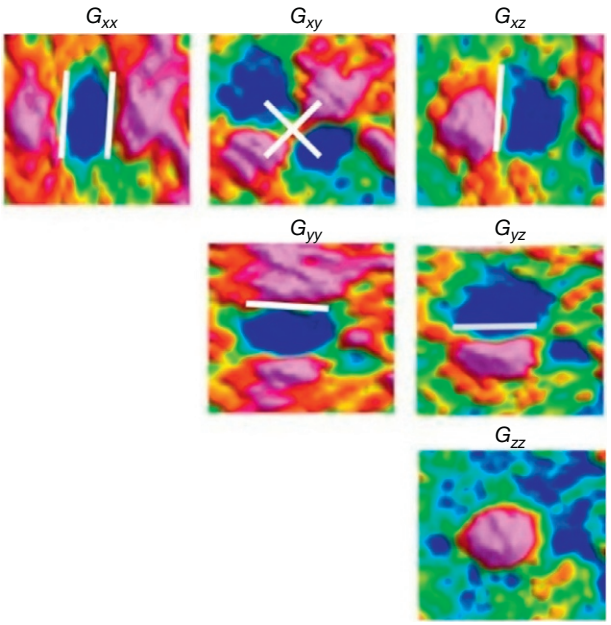


Figure 35 The response of each independent tensor component across a body with a positive density contrast. The gravity values are only relative with the reds being gravity gradient high values and the blues being gravity gradient low values. G_{zz} locates 'target'; G_{xx} and G_{yy} identify north–south and east–west edges of target, respectively; G_{xz} and G_{yz} identify central axes of target mass with highs and lows defining fault trends; and G_{xy} shows four anomalies that point to the center of mass. Modified from Mims JH and Mataragio J (2010) Airborne full tensor gravity. *Proceedings, Symposium on the Application of Geophysics to Environmental and Engineering Problems, Keystone, Colorado*, pp. 438–442.

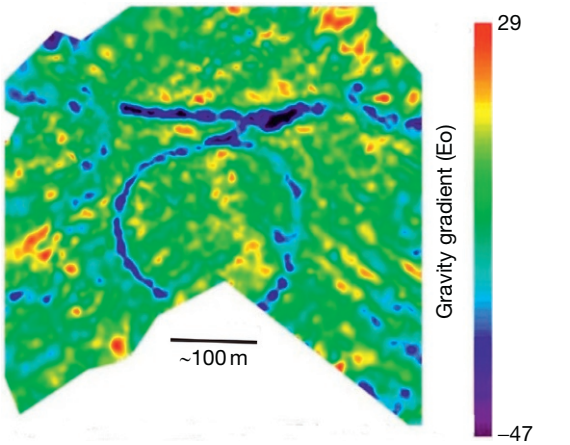


Figure 36 Helicopter Falcon airborne gravity gradiometer survey flown over the tunnel for the Superconducting Super Collider. The plot shows the terrain-corrected G_{zz} component. Modified from Hodges G, Dransfield M, and Shei T-C (2010) The Falcon airborne gravity gradiometer for engineering applications. *Proceedings, Symposium on the Application of Geophysics to Environmental and Engineering Problems, Keystone, Colorado*, pp. 443–447.

contrast. The three diagonal components are shown even though only two of them are independent. A good source of information on gravity gradiometry is the Australian Society of Exploration Geophysicists' Airborne Gravity and Gradiometry 2004 Workshop, which took place in Sydney, Australia, in 2004. Papers by Pawlowski (1998), Hodges et al. (2010), Mims and Mataragio (2010), Barnes and Lumley (2011), Martinez et al. (2013), and **Chapter 11.12** also provide useful examples on gravity gradiometry. **Figure 3** in the paper by van Leeuwen (2000) provides a nice comparison of the computed ground G_{zz} component with the measured airborne G_{zz} component flown at an altitude of 120 m. Both the data sets were terrain-corrected. **Figure 36** from Hodges et al. (2010) provides an example of the amazing resolution of the helicopter Falcon gravity gradiometer system. The system was flown at very low altitude (<10 m) over the tunnel for the Superconducting Super Collider (a circular ring more than 100 m in diameter with dimensions 4×4 m and buried 5 m below the surface). The ring shows up as a gravity gradient low as expected (**Figure 36**).

11.15.6.2 Land-Based Seismic Methods in Mineral Exploration

Seismic methods were not used for hard-rock mineral exploration until relatively recently. However, they were applied to mineral exploration in sedimentary environments because of the similarities to petroleum exploration (e.g., uranium exploration in the Athabasca Basin (**Section 11.15.4.2.4** and O'Dowd et al., 2006) and potash exploration in Saskatchewan (**Figure 32** in **Section 11.15.6**)). Advances in hard-rock acquisition and processing procedures plus more and better measurements of density and velocity of igneous and metamorphic rocks and minerals provided the impetus for the development of 2-D and 3-D seismic methods for hard-rock mining applications. For example, the number of geophone channels has significantly increased so that now, 10 000 or more channels can be deployed. This leads to better spatial sampling, sampling larger offsets, and the collection of multicomponent seismic information. Acquisition systems can now be deployed without using cables (wireless acquisition systems) making the collection of data in difficult areas easier. See the special section entitled 'Seismic methods in mineral exploration' in Geophysics (2012) (issue 5, volume 77) and Eaton et al. (2003) for recent information on seismic applications to mineral exploration. Reflection seismic methods are also used in conjunction with other information for mine planning (Malehmir et al., 2012a,b).

The following provides an example of 3-D seismic imaging of VMS deposits in the Flin Flon mining camp in Manitoba, Canada (Malinowski et al., 2012; White et al., 2012). The Flin Flon mining camp in Manitoba is located within the Paleoproterozoic Flin Flon metavolcanic belt and is close to the town of Flin Flon, Manitoba, near the Saskatchewan border at ~55° north. The Main, 777, and Callinan deposits lie along the west limb of the Hidden Lake synform (White et al., 2012). The Flin Flon Main deposit closed in 1990 but the 777 (14.2 million tons) and Callinan (8.4 million tons) deposits are still being mined. However, the Callinan deposit was nearly mined out when the 3-D survey was carried out in 2007. The 777 deposits are generally more massive than the Callinan deposits and contain more pyrrhotite than pyrite relative to the Callinan deposits (White et al., 2012). The deposits tend to be elongated in the dip direction and have thicknesses between 10 and 30 m (**Figure 37**).

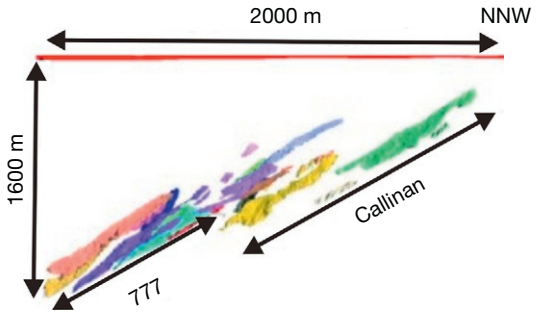


Figure 37 Cross section of the Callinan and 777 deposits (individual deposits in different colors) showing them plunging to the SSE. Modified from White DJ, Secord D, and Malinowski M (2012) 3D seismic imaging of volcanogenic massive sulfide deposits in the Flin Flon mining camp, Canada: Part 1 – Seismic results. *Geophysics* 77: WC47–WC58.

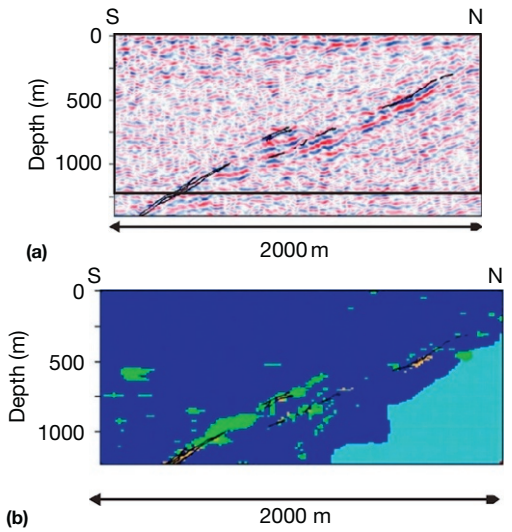


Figure 38 (a) Vertical slice through the 3-D migrated volume. The black rectangle outlines the area covered by the model in (b). (b) Corresponding vertical slice from the 3-D lithologic model based on drillhole information and used for seismic modeling (see Malinowski et al. (2012) for more details on the modeling). Dark blue, basalt; light blue, sandstone; green, rhyolite; and orange, ore. The ore lenses are shown in black in the two plots. Modified from White DJ, Secord D, and Malinowski M (2012) 3D seismic imaging of volcanogenic massive sulfide deposits in the Flin Flon mining camp, Canada: Part 1 – Seismic results. *Geophysics* 77: WC47–WC58.

The 3-D seismic survey covered an area of $\sim$17 km^2 with over 900 shot locations and more than 3300 receiver stations. For details of the processing procedures, see White et al. (2012). Refraction and residual static corrections were carried out along with a conventional dip-moveout (DMO) time processing flow. Poststack migration, poststack DMO migration, and prestack time migration were carried out on the data volume. The poststack DMO migration visually provided the best results.

Figure 38(a) shows the migrated data for a vertical section over the Callinan and 777 ore zones. The black lines are the locations of the ore lenses. Note that the correlation with reflectors is better for the Callinan zones (**Figure 38(b)**). More diffractions occur at depth for the 777 ore lenses. Further details on the processing and interpretation of these data can be found in White et al. (2012).

11.15.7 Future Developments

This chapter provides a snapshot into the present 'state-of-the-art' application of geophysical techniques to mineral exploration, development, and reclamation. New methods have been introduced to explore deeper and to provide higher spatial resolution. Inversion schemes are now available to provide for better modeling of the physical parameters of rocks and minerals. We have seen some mineral deposit types require the use of specific geophysical techniques or combinations of geophysical techniques to effectively explore for them. What new developments can we expect to see in the future to further enhance our ability to explore and develop mineral resources?

Future geophysical developments are difficult to predict, although there are trends that can be observed within the geophysical community. We can expect to see inversions of large, regional data sets (magnetic, electromagnetic, radiometric, etc.) to become commonplace in the future (Cox and Zhdanov, 2008; Cox et al., 2010; Wilson et al., 2010; Zhdanov et al., 2011). Refinement of techniques to perform joint inversions of several geophysical data sets simultaneously will be developed.

There is significant interest in integrating drone (unmanned aircraft) technology and geophysical surveying (Miles et al., 2008). Recent development of light weight drones will impact how local, small-scale geophysical surveys are carried out in the future over specific mineral claim blocks. Airborne gravity will continue to develop both with platform design and new algorithms to utilize the vast amount of gradient data generated from the surveys.

Rapid, low-cost geophysical methods are being developed to reduce the environmental impact (such as reduction in line cutting, use of smaller field crews, and use of lighter and more portable equipment) of exploration. Cost-effective 3-D acquisition systems for resistivity and IP data will be developed to provide enhanced coverage and interpretation when exploring for porphyry and gold deposits, for example.

Further application of 3-D seismic methods will be pursued, and 3-D seismic will become a standard tool used for mine development. Further refinements in the interpretation of high-velocity volcanic and plutonic rocks will be made, and research will continue into the relationship between rock properties and seismic interpretation for hard-rock environments.

References

Aleva GJJ (Compiler) (1994) *Laterites: Concepts, Geology, Morphology and Chemistry*, 169 pp. Wageningen: ISRIC, ISBN: 90-6672-053-0.

Annan AP, Davis JL, and Gendzwill D (1988) Radar sounding in potash mines, Saskatchewan, Canada. *Geophysics* 53: 1556–1564.

Argyle M, Ferguson S, Sander L, and Sander S (2000) AIRGrav results: A comparison of airborne gravity data with GSC test site data, Ottawa, Canada. *The Leading Edge* October: 1134–1138.

Balch SE (1999) Ni–Cu sulphide deposits with examples from Voisey's Bay. In: Lowe C, Thomas MD, and Morris WA (eds.) *Geophysics in Mineral Exploration:*

Fundamentals and Case Histories. *Short Course Notes*, vol. 14, pp. 21–40. Newfoundland: Geological Association of Canada, Memorial University.
Barnes G and Lumley J (2011) Processing gravity gradient data. *Geophysics* 76: I33–I47.
Barrie CT and Hannington MD (eds.) (1999a) *Reviews in Economic Geology*, vol. 8. *Volcanic-Associated Massive Sulfide Deposits: Processes and Examples in Modern and Ancient Settings*, p. 408. Denver: Society of Economic Geologists.
Barrie CT and Hannington MD (1999b) Classification of volcanic-associated massive sulfide deposits based on host-rock composition. *Reviews in Economic Geology* 8: 1–11.
Bell K (1989) *Editor, Carbonatites: Genesis and Evolution*. London: Unwin Hyman.
Bergman SC (1987) Lamproites and other potassium-rich igneous rocks: A review of their occurrences, mineralogy and geochemistry. In: Fitton JG and Upton BGJ (eds.) *Alkaline Igneous Rocks*, pp. 103–119. London: Geological Society of London, Special Publication No. 30.
Best ME, Abercrombie HJ, and Peirce JW (1998) Interpreted faulting patterns in northeast Alberta using high resolution aeromagnetic data. *Canadian Journal of Exploration Geophysics* 34: 48–57.
Boivin M (2007) Advances in geophysical technology for VMS exploration. In: Milkereit B (ed.) *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, pp. 731–739.
Bower B, Payne J, DeLong C, and Rebagliati CM (1995) The oxide-gold, supergene and hypogene zones at the Casino gold–copper–molybdenum deposit, west central Yukon. In: Schroeter TG (ed.) *Detecting Ore Using GRS and K Alteration, Porphyry Deposits of the Northwestern Cordillera of North America*, Special Volume 46, pp. 352–366. Montreal, Quebec: Canadian Institute of Mining, Metallurgy, and Petroleum.
Brant AA (1966) Geophysics in the exploration for Arizona porphyry deposits. In: Titley SR and Hicks CL (eds.) *Geology of the Porphyry Copper Deposits, Southwestern North America*, pp. 87–110. Tucson, AZ: University of Arizona Press.
Collett LS (1986) Development of the airborne electromagnetic technique. In: Palacky GJ (ed.) *Airborne Resistivity Mapping*, pp. 9–18. Ottawa: Geological Survey of Canada Paper 66-2.
Cooper NJ and Swift R (1994) The application of TEM to Cyprus-type massive sulfide exploration in Cyprus. *Geophysics* 59: 202–214.
Corry CE, Emer DF, and Zonge KL (1988) Controlled source audiofrequency magnetotelluric surveys of porphyry sulfide deposits and prospects in the cordillera of the United States. In: *North American Conference on Tectonic Control of Ore Deposits and the Vertical and Horizontal Extent of Ore Systems, University of Missouri, Rolla, Proceedings Volume*, pp. 204–213.
Cox LH and Zhdanov MS (2008) Advanced computational methods of rapid and rigorous 3-D inversion of airborne electromagnetic data. *Communications in Computational Physics* 3: 160–179.
Cox LH, Wilson GA, and Zhdanov MS (2010) 3D inversion of airborne electromagnetic data using a moving footprint. *Exploration Geophysics* 41: 250–259.
Davis JL and Annan AP (1989) Ground penetrating radar for high-resolution mapping of soil and rock properties. *Geophysics, Prospecting* 37: 531–551.
Davis EE, Goodfellow WD, Bornhold BD, et al. (1987) Massive sulfides in a sedimented rift valley, northern Juan de Fuca Ridge. *Earth and Planetary Science Letters* 82: 49–61.
Dobrin MB (1960) *Introduction to Geophysical Prospecting*, 2nd edn. New York: McGraw-Hill Book Company.
Dransfield MH and Lee JB (2004) The FALCON® airborne gravity gradiometer systems. In: Lane RJL (ed.) *Airborne Gravity 2004 – Abstracts from the ASEG-PESA Airborne Gravity 2004 Workshop, Geoscience Australia Record 2004/18*, pp. 15–20.
Duckworth RC, Fallick AE, and Rickard D (1994) Minerology and sulfer isotope composition of thw Middle Valley massive sulfide deposit: Northern Juan de Fuca rudge. In: Mottl MJ, Davis EE, Fisher AT, and Slack JF (eds.) *Scientific Results*, vol. 139, *Proceedings of the Ocean Drilling Program*.
Eaton DW, Milkereit B, and Salisbury MH (eds.) (2003) *Geophysical Development Series*, vol. 10. *Hardrock Seismic Exploration*. Tulsa, OK: Society of Exploration Geophysicists.
Eckstrand OR and Hulbert LJ (2007) Magmatic nickel–copper–platinum group element deposits. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 205–222. Newfoundland: Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5, Memorial University.
Einaudi MT, Meinert LD, and Newberry RJ (1981) Skarn deposits. *Economic Geology*, 75th Anniversary Volume, pp. 317–391. Denver: Society of Economic Geologists.
Emsbo P (2009) Geologic criteria for the assessment of sedimentary exhalative (sedex) Zn-Pb-Ag deposits. *U.S. Geological Survey Open-File Report 2009-1209*, p. 21.
Ford K, Keating P, and Thomas MD (2007) Overview of geophysical signatures associated with Canadian ore deposits. In: *Mineral Resources of Canada: A Synthesis of Major Deposit-types*, p. 21, Geological Survey of Canada.
Forsberg R and Olsen AV (2010) Airborne gravity field determination. In: Xu G (ed.) *Sciences of Geodesy*, pp. 83–104. Berlin Heidelberg: Springer.
Foss C (2004) Precision in location of vertical pipes from inversion of airborne gravity gradient data. In: *Society of Exploration Geophysics 74th Annual Meeting, Denver, Colorado*.
Francke J (2012) A review of selected ground penetrating radar applications to mineral resource evaluations. *Journal of Applied Geophysics* 81: 29–37.
Franklin JM (1996) Volcanic-associated massive sulfide base metals. In: Eckstrand OR, Sinclair WD, and Thorpe RI (eds.) *Geology of Canadian Mineral Deposit Types*, vol. 8, pp. 158–183. Ottawa: Geology of Canada, Geological Survey of Canada.
Franklin JM, Sangster DM, and Lydon JW (1981) Volcanic-associated massive sulfide deposits. In: Skinner BJ (ed.) *Economic Geology,* 75th Anniversary Volume, pp. 485–627. Denver: Society of Economic Geologists.
Franklin JM, Gibson HL, Galley AG, and Jonasson IR (2005) Volcanogenic massive sulfide deposits. In: Hedenquist JW, Thompson JFH, Goldfarb RJ, and Richards JP (eds.) *Economic Geology,* 100th Anniversary Volume, pp. 523–560. Littleton, CO: Society of Economic Geologists.
Fullagar P (1989) Generation of conductivity-depth pseudo-sections for coincident loop and in-loop TEM data. In: *7th Geophysical Conference, Australian Society of Exploration Geophysicists*, vol. 20, pp. 43–45.
Gasperikova E, Cuevas NH, and Morrison HF (2005) Natural field induced polarization for mapping of deep mineral deposits: A field example from Arizona. *Geophysics* 70: B61–B66.
Gendzwill DJ and Brehm R (1993) High-resolution seismic reflections in a potash mine. *Geophysics* 58: 741–748.
Golightly JP (1981) Nickeliferous laterite deposits. *Economic Geology* 75: 710–735.
Goodfellow WD, McCutcheon SR, and Peter JM (2003) Massive sulfide deposits of the Bathurst Mining Camp, New Brunswick and Northern Maine: Introduction and summary of findings. *Economic Geology Monograph* 11: 1–16.
Haijnal A, Pandit B, Takacs E, Annesley I, and Wallster D (2007) Recent development in 2D and 3D seismic imaging of high-grade uranium ore deposit related environments. In: Milkereit B (ed.) *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, pp. 1131–1135.
Hammond S (2003) Air-FTG(tm): Bell geospace's airborne gravity gradiometer – A description and case study. *Preview:* 24–26.
Hattula A and Rekola T (2000) Exploration geophysics at the Pyhasalmi mine and grade control work of the Outokumpu Group. *Geophysics* 65: 1961–1969.
Hearst RB, Morris WA, and Schieck DG (1994) Reflection seismic profiling for massive sulfides in the high Arctic. In: *64th Annual International Meeting, Society of Exploration Geophysicists*, pp. 531–533.
Hodges G, Dransfield M, and Shei T-C (2010) The Falcon airborne gravity gradiometer for engineering applications. In: *Proceedings, Symposium on the Application of Geophysics to Environmental and Engineering Problems, Keystone, Colorado*, pp. 443–447.
Hoeve J and Quirt DH (1984) Mineralization and host rock alteration in relation to clay mineral diagenesis and evolution of the Middle-Proterozoic, Athabasca Basin, northern Saskatchewan, Canada. *Saskatchewan Research Council*, *SRC Technical Report 187*, p. 187.
Hoffman PF (1990) Subdivision of the Churchill Province and extent of the Trans-Hudson Orogen. In: Lewry JF and Stauffer MR (eds.) *The Early-Proterozoic Trans-Hudson Orogen of North America*, pp. 15–39. Newfoundland: Geological Association of Canada, Special Paper 37, Memorial University.
Hoover DB, Heran WD, and Hill PL (1992) The geophysical expression of selected mineral deposit models. *United States Department of the Interior*, *Geological Survey*, *Open File Report 92-557*.
Huang H and Rudd J (2006) Conductivity-depth imaging of time-domain EM data based on pseudo-layer half-space model. In: *Society of Exploration Geophysics Annual Meeting, New Orleans, USA*.
Interpex (2006) *IX1D.v3 Tutorials*. Golden, CO: Interpex Limited.
Irvine RJ and Smith MJ (1990) Geophysical exploration for epithermal gold deposits. *Journal of Geochemical Exploration* 36: 375–412.
Jansen J and Witherly K (2004) The Tli Kwi Cho Kimberlite Complex, Northwest Territories, Canada: A geophysical case study. In: *Society of Exploration Geophysicists 74th Annual Meeting, Denver, Colorado*.
Jefferson CW and Delaney G (eds.) (2007) *Geology and Exploration Technology of the Proterozoic Athabasca Basin, Saskatchewan and Alberta*, p. 644. Mineral Deposits Division, Geological Association of Canada; Memorial University, Geological Survey of Canada, Ottawa Bulletin 588; Saskatchewan Geological Survey Special Publication 18, Special Publication 4, Saskatoon, Saskatchewan.

Johnson IM and Fujita M (1985) The Hishikari gold discovery: An airborne EM discovery. *Bulletin of the Canadian Institute of Mining and Metallurgy* 78: 61–66.
Juhojuntti N, Wood G, Juhlin C, O'Dowd C, Dueck P, and Cosma C (2012) 3D seismic survey at the Millennium uranium deposit, Saskatchewan, Canada: Mapping depth to basement and imaging post-Athabasca structure near the orebody. *Geophysics* 77: WC245–WC258.
Kawasaki TW, Okada K, and Kubota R (1986) Geophysical surveys in the Hishikari mine area. *Mining Geology* 36: 131–147.
Kearey P and Brooks M (1984) *An Introduction to Geophysical Exploration*, p. 296. Oxford, England: Blackwell Scientific.
Keating P and Sailhac P (2004) Use of the analytic signal to identify magnetic anomalies due to kimberlite pipes. *Geophysics* 69: 180–190.
Keller GV (1988) Rock and mineral properties. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics*, vol. 1. Tulsa, OK: Society of Exploration Geophysicists, Investigations in Geophysics No. 3.
Keller GV and Frischknecht FC (1966) *Electrical Methods in Geophysical Prospecting*, p. 517. London: Pergamon Press.
Killeen PG (1979) Gamma-ray spectrometric methods in uranium exploration – Application and interpretation. In: Hood PJ (ed.) *Geophysics and Geochemistry in Search for Metallic Ores*, pp. 163–230. Ottawa: Geological Survey of Canada, Economic Geology Report 31.
Kirkham RV (1971) Intermineral intrusions and their bearing on the origin of porphyry copper and molybdenum deposits. *Economic Geology* 66: 1244–1250.
Kirkham RV (1972) Porphyry deposits. In: Blackadar RG (ed.) *Report of Activities Part B, November 1971 to March 1972*, pp. 62–64, Geological Survey of Canada, Paper 72-1b.
Kirkham RV (1984) Sedimentary copper. In: Eckstrand OR (ed.) *Canadian Mineral Deposit Types: A Geological Synopsis*, p. 27, Geological Survey of Canada Economic Geology Report 36.
Kirkham RV (1989) Distribution, settings, and genesis of sediment-hosted stratiform copper deposits. In: Boyle RW, Brown AC, Jefferson CW, Jowett EC, and Kirkham RV (eds.) *Sediment-Hosted Stratiform Copper Deposits*, pp. 3–38. Newfoundland: Geological Association of Canada, Special Paper 36, Memorial University.
Kirkham RV and Sinclair WD (1996) Porphyry copper, gold, molybdenum, tungsten, tin, silver. In: Eckstrand OR, Sinclair WD, and Thorpe RI (eds.) *Geology of Canadian Mineral Deposit Types*, pp. 421–446. Geology of Canada, Geological Survey of Canada, No. 8.
Kjarsgaard BA (2007) Kimberlite pipe models: Significance for exploration. In: Milkereit B (ed.) *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, pp. 667–677.
Kovin ON (2010) Fracture imaging in the upper Kama potash mine using 3D GPR data. In: *Expanded Abstract, Symposium on the Application of Geophysics to Environmental and Engineering Problems, Keystone, Colorado*.
Kresten P (1983) Carbonatite nomenclature. *International Journal of Earth Sciences* 72.
Lajoie JL and Klein J (1979) Geophysical exploration at the Pine Point Mines Ltd. zinc-lead property, Northwest Territories, Canada. In: Hood PJ (ed.) *Geophysics and Geochemistry in the Search for Metallic Ores*, *Proceedings of Exploration 1977, Geological Survey of Canada, Economic Geology Report 31*, pp. 652–664.
Lee JB (2001) FALCON gravity gradiometer technology. *Exploration Geophysics* 32: 247–250.
Legault J, Carriere D, and Petrie L (2008) Synthetic model testing and distributed acquisition dc resistivity results over an unconformity uranium target from the Athabasca Basin, northern Saskatchewan. *The Leading Edge:* 46–51.
Li Y and Oldenburg DW (1996) 3-D inversion of magnetic data. *Geophysics* 61: 394–408.
Li Y and Oldenburg DW (1998) 3D inversion of gravity data. *Geophysics* 63: 109–119.
Lydon JW (1984) Ore deposit models. 8. Volcanogenic sulfide deposits: Part I. A descriptive model. *Geoscience Canada* 11: 195–202.
Lowe C, Thomas MD, and Morris WA (1999) In: *Geophysics in Mineral Exploration: Fundamentals and Case Histories*. *Short Course Notes*, vol. 14, p. 175. Newfoundland: Geological Association of Canada, Memorial University.
Lydon JW (1995) Sedimentary exhalative sulfides (SEDEX). In: Eckstrand OR, Sinclair WD, and Thorpe RI (eds.) *Geology of Canadian Mineral Deposit Types*, pp. 30–152. Geology of Canada, Geological Survey of Canada, No. 8.
Lydon JW, Goodfellow WD, Dubé B, et al. (2004) A preliminary overview of Canada's mineral resources. *Geological Survey of Canada Open File 4668*, p. 20, Ottawa.
Lindsay DH, Andreasen GE, and Balsley JR (1966) Magnetic properties of rocks and minerals. In: *Handbook of Physical Constants*. revised edn., *Memoir*, revised edn., vol. 95, p. 543. New York: Geological Society of America, Section 25, Page 46 of 46.
Loke MH and Barker RD (1996a) Rapid least-squares inversion of apparent resistivity pseudosections using a quasi-Newton method. *Geophysical Prospecting* 44: 131–152.
Loke MH and Barker RD (1996b) Practical techniques for 3D resistivity surveys and data inversion. *Geophysical Prospecting* 44: 499–523.
Macnae JC (1979) Kimberlites and exploration geophysics. *Geophysics* 44: 1395–1416.
Macnae JC, Smith R, Polzer BD, Lamontagne Y, and Klinkert PS (1991) Conductivity-depth imaging of airborne electromagnetic step-response data. *Geophysics* 56: 102–114.
Malehmir A, Juhlin C, Wijns C, Urosevic M, Valasti P, and Koivisto E (2012a) 3D reflection seismic imaging for open-pit mine planning and deep exploration in the Kevitsa Ni-Cu-PGE deposit, northern Finland. *Geophysics* 77: WC95–WC108.
Malehmir A, Durrheim R, Bellefleur G, et al. (2012b) Seismic methods in mineral exploration and mine planning: A general overview of past and present case histories and a look into the future. *Geophysics* 77: WC173–WC190.
Malinowski M, Schetselaar E, and White DJ (2012) 3D seismic imaging of volcanogenic massive sulfide deposits in the Flin Flon mining camp, Canada: Part 2 – Forward modeling. *Geophysics* 77: WC81–WC93.
Martinez C, Li Y, Krahenbuhl R, and Braga AB (2013) 3D inversion of airborne gravity gradiometry data in mineral exploration: A case study in the Quadrilátero Ferrífero, Brazil. *Geophysics* 78: B1–B11.
Meinert LD (1982) Skarn, manto, and breccia pipe formation in sedimentary rocks of the Cananea mining district, Sonora, Mexico. *Economic Geology* 77: 919–949.
Meissner R (2002) *The Little Book of Planet Earth*, p. 202. New York: Copernicus Books.
Melvin JL (ed.) (1991) *Evaporites, Petroleum and Mineral Resources*. Amsterdam: Elsevier.
Miller RD, Xia J, Park CB, and Ivanov JM (1999) Multichannel analysis of surface waves to map bedrock, Kansas Geological Survey. *The Leading Edge* December: 1392–1396.
Mims JH and Mataragio J (2010) Airborne full tensor gravity. In: *Proceedings, Symposium on the Application of Geophysics to Environmental and Engineering Problems, Keystone, Colorado*, pp. 438–442.
Mitchell RH (1986) *Kimberlites: Mineralogy, Geochemistry, and Petrography*, p. 442. New York: Plenum Press.
Miles PJ, Partner RT, Keeler KR, and McConnell TJ (2008) Unmanned airborne vehicle for geophysical surveying. *US Patent US2008/0125920 AI*.
Morgan LA (2012) Geophysical characteristics of volcanogenic massive sulfide deposits in volcanogenic massive sulfide occurrence model. *U.S. Geological Survey Scientific Investigations Report 2010-5070-C*, Chapter 7, p. 16.
Morison SR (1989) Placer deposits in Canada. In: Fulton RJ (ed.) *Quaternary Geology of Canada and Greenland*, pp. 687–694. Ottawa: Geology of Canada, Geological Survey of Canada.
Morris B, Leblanc G, and Prevec I (2002) An alternative simple procedure to identify magnetic and other geophysical anomalies due to kimberlite pipes. In: *72nd Annual Meeting, Society of Exploration Geophysics, Salt Lake City, Utah*.
Munday T, Sumpton J, and Fitzpatrick A (2004) Exploration for kimberlites through a complex regolith cover – A case study in the application of AEM in the deeply weathered Archaean Yilgarn Craton, Western Australia. In: *74th Annual Meeting, Society of Exploration Geophysicists, Denver, Colorado*.
Nabighian MN and Asten MW (2002) Metalliferous mining geophysics—State of the art in the last decade of the 20th century and the beginning of the new millennium. *Geophysics* 67: 964–978.
Naldrett AJ (2004) *Magmatic Sulfide Deposits—Geology, Geochemistry, and Exploration*. Berlin: Springer, p. 727.
Nicol A and Lawton DC (2012) Multi-component investigation of fracturing in a potash mining region in Saskatchewan, Canada. In: *Expanded abstracts, Society of Exploration Geophysics Annual Meeting, Las Vegas, Nevada*.
Nimick G and Koch R (2008) A progressive geophysical exploration strategy at the Shea Creek uranium deposit. *The Leading Edge* January: 52–63.
O'Dowd C, Wood G, Brisbin D, and Powell B (2006) Enhancing uranium exploration through seismic methods and potential field modeling at the McArthur River mine site, Saskatchewan, Canada. In: *Society of Exploration Geophysicists Annual Meeting, New Orleans*, pp. 1253–1257.
Palacky GJ (1988) Resistivity characteristics of geologic targets. In: Nabighian MN (ed.) *Electromagnetic Methods in Applied Geophysics*, vol. 1 Tulsa, OK: Society of Exploration Geophysicists, Investigations in Geophysics No. 3.
Parasinis DS (1975) *Mining Geophysics*. Amsterdam: Elsevier.
Pare P and Legault JM (2010) Ground IP-resistivity, and airborne spectrem and helicopter ZTEM survey results over Pebble copper-moly-gold porphyry deposit, Alaska. In: *Society of Exploration Geophysicists Annual Meeting, Denver, Colorado*.
Park CB, Miller RD, and Xia J (1999) Multichannel analysis of surface waves. *Geophysics* 64: 800–808.
Pawlowski B (1998) Gravity gradiometry in resource exploration. *The Leading Edge* January: 51–52.

Pesowski MS and Larson RK (2000) Seismic exploration methods applied to potash mining: Risk analysis and mine planning. In: Expanded Abstracts. Tulsa, OK: Society of Exploration Geophysics.

Poulsen KH and Hannington MD (1996) Volcanic-associated massive sulfide gold. In: Eckstrand OR, Sinclair WD, and Thorpe RI (eds.) *Geology of Canadian Mineral Deposit Types*, vol. 8, pp. 183–196. Ottawa: Geology of Canada, Geological Survey of Canada.

Powell B, Wood G, and Bzdel L (2007) Advances in geophysical exploration for uranium deposits in the Athabasca Basin. In: Milkereit B (ed.) *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration,* pp. 771–790.

Pringle JK, Styles P, Howell CP, Branston MW, Furner R, and Toon SM (2012) Long-term time-lapse microgravity and geotechnical monitoring of relict salt mines, Marston, Cheshire, UK. *Geophysics* 77: B287–B294.

Rebagliati CM, Lang J, Titley E, et al. (2008) *Technical Report on the 2008 Program and Update on Mineral Resources and Metallurgy*, *Pebble Copper-Gold-Molybdenum Project*, *Iliamna Lake area*, *Southwestern Alaska*, *USA*, *for Northern Dynasty Minerals Ltd*., *NI-43 101 Report*, 170 pp.

Reitz JR and Milford FJ (1962) *Foundations of Electromagnetic Theory*, p. 387. Reading, MA: Addison Wesley Publishing Company.

Robertson K and Megraw P (2009) Cinco de Mayo: A new silver, lead, and zinc discovery in northern Mexico. *The Leading Edge* June: 730–735.

Rona PA (1978) Magnetic signature of hydrothermal alteration and volcanic mineral deposits in oceanic crust. *Journal of Volcanology and Geothermal Research* 3: 219–225.

Ruotoistenmäki T (2008) Geophysical maps and petrophysical data of Mozambique. *Geological Survey of Finland*, *Special Paper 48*, pp. 65–80.

Ruzicka V (1995a) Unconformity-associated uranium. In: Eckstrand OR, Sinclair WD, and Thorpe RI (eds.) *Geology of Canadian Mineral Deposit Types*, pp. 97–210. Ottawa: Geology of Canada, Geological Survey of Canada, No. 8.

Ruzicka V (1995b) Unconformity type uranium deposits. In: Kirkham RV, Sinclair WD, Thorpe RT, and Duke JM (eds.) *Mineral Deposit Modeling*, pp. 125–150. Newfoundland: Geological Association of Canada, Special Paper 40, Memorial University.

Sander S, Argyle M, Elieff S, Fergusun S, Lavoi V, and Sander L (2005) The AIRGrav airborne gravity system. *Canadian Society of Exploration Geophysics Recorder* October: 5.

Sangster DF (1995) Mississippi Valley-type lead zinc. In: Eckstrand OR, Sinclair WD, and Thorpe RI (eds.) *Geology of Canadian Mineral Deposit Types*, pp. 253–261. Ottawa: Geology of Canada, Geological Survey of Canada, No. 8.

Salisbury MH, Harvey CW, and Mathews L (2003) The acoustic properties of ores and host rocks in hardrock terranes. In: Eaton DW, Milkereit B, and Salisbury MH ,(eds.) *Hardrock Seismic Exploration*. Tulsa, OK: Society of Exploration Geophysics, Geophysical Development No. 10.

Seigel HO, Hill HL, and Baird JG (1968) Discovery case history of the Pyramid ore bodies Pine Point, Northwest Territories, Canada. *Geophysics* 33: 645–656.

Seigel H, Nabighian M, Parasnis D, and Vozoff K (2007) The early history of the induced polarization method. *The Leading Edge* March: 312–321.

Shives RBK, Charbonneau BE, and Ford KL (2000) The detection of potassic alteration by gamma-ray spectrometry—Recognition of alteration related to mineralization. *Geophysics* 65: 2001–2011.

Shuey RT (1975) *Semiconducting Ore Minerals*, p. 415. Amsterdam: Elsevier Scientific Publishing Company.

Sinclair WD (2007) Porphyry deposits. In: Goodfellow WD (ed.) *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods*, pp. 223–243. Newfoundland: Geological Association of Canada, Mineral Deposits Division, Special Publication No. 5, Memorial University.

Skirrow RG, Jaireth S, and Huston DL, et al. (2009) Uranium mineral systems: Processes, exploration criteria and a new deposit framework. *Geoscience Australia Record 2009/20*, p. 44.

Smit JP and Roux Le (2009) TDEM survey at Shea Creek uranium deposit utilizing a low-temperature superconducting SQUID. In: *11th Annual South African Geophysical Association Biennial Technical Meeting and Exhibition, Swaziland,* pp. 46–49.

Smith RS, Annan AP, Lemieux J, and Pedersen RN (1996) Application of a modified GEOTEM system to reconnaissance exploration for kimberlites in the Point Lake area, NWT, Canada. *Geophysics* 61: 82–92.

Sumner JS (1976) *Principles of Induced Polarization for Geophysical Exploration*, p. 277. New York: Elsevier.

Sutherland DG (ed.) (1991) *Alluvial Mining*, p. 601. London: Institution of Mining and Metallurgy, Elsevier Applied Science.

Telford WM, Geldart LP, Sherriff RE, and Keys DA (1976) *Applied Geophysics*. London: Cambridge University Press.

Tixier MP and Alger RP (1970) Log evaluation of nonmetallic mineral deposits. *Geophysics* 35: 124–142.

Tuncer V, Unsworth MJ, Weerachai S, and Craven JA (2006) Exploration for unconformity-type uranium deposits with audiomagnetotelluric data: A case study from the McArthur River mine, Saskatchewan, Canada. *Geophysics* 71: B201–B209.

Vallee MA, Smith RS, and Keating P (2011) Metalliferous mining geophysics—State of the art after a decade in the new millennium. *Geophysics* 76: W31–W50.

van Leeuwen E (2000) BHP develops airborne gravity gradiometer for mineral exploration. *The Leading Edge* December: 1296–1297.

Vasanthi A and Mallick K (2005) Bouguer gravity anomalies and occurrence patterns of kimberlite pipes in Narayanpet-Maddur Regions Andhra Pradesh, India. *Geophysics* 70: J13–J24.

Verdun J, Bayer R, Klingele EE, Cocard M, Geiger A, and Halliday M (2002) Airborne gravity measurements over mountainous areas by using a LaCoste & Romberg air-sea gravity meter. *Geophysics* 67(3): 807–816.

Wegener A (1966) *The Origin of Continents and Oceans*, p. 246. Mineola, NY: Courier Dover, Biram John (translator).

White DJ, Secord D, and Malinowski M (2012) 3D seismic imaging of volcanogenic massive sulfide deposits in the Flin Flon mining camp, Canada: Part 1—Seismic results. *Geophysics* 77: WC47–WC58.

Williams PK (1997) Towards a multidisciplinary integrated exploration process for gold discovery. In: Gubins AG (ed.) *Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral Exploration,* pp. 1015–1028.

Wilson GA, Cox LH, and Zhdanov MS (2010) Practical 3D inversion of entire airborne electromagnetic surveys. *Preview* 146: 29–33.

Xue G-Q, Qin K-Z, Xiu L, Li G-M, Qi Z-P, and Zhou N-N (2012) Discovery of a large-scale porphyry molybdenum deposit in Tibet through a modified tem exploration method. *Journal of Environmental and Engineering Geophysics* 17: 19–25.

Yang D and Oldenburg DW (2012) Three-dimensional inversion of airborne time-domain electromagnetic data with applications to a porphyry deposit. *Geophysics* 77: B23–B34.

Zhdanov MS, Smith RB, Gribenko A, Cuma M, and Green M (2011) Three dimensional inversion of large scale earthscope magnetotelluric data based on the integral equation method: Geoelectrical imaging of the Yellowstone conductive mantle plume. *Geophysical Research Letters* 38: L08307, 7 pp.

11.16 Unconventional Fossil Fuel Reservoirs and Water Resources

E Barth-Naftilan and JE Saiers, Yale School of Forestry and Environmental Studies, New Haven, CT, USA

© 2015 Elsevier B.V. All rights reserved.

11.16.1	**Introduction**	557
11.16.2	**Unconventional Resources**	558
11.16.3	**Shale Gas and Shale Oil**	560
11.16.3.1	Shale Gas and Shale Oil Extraction	560
11.16.4	**Heavy Oil and Oil Sands**	561
11.16.4.1	Heavy Oil and Oil Sands Extraction	562
11.16.5	**Water Issues Associated with Unconventional Resource Extraction**	562
11.16.5.1	Water Resources and Shale Gas/Shale Oil Extraction	563
11.16.5.1.1	Risk of water contamination with frac fluids and brines	563
11.16.5.1.2	Risk of water contamination with methane	564
11.16.5.1.3	Water resource appropriation and shale gas/shale oil extraction	565
11.16.5.2	Water Resources and Heavy Oil/Oil Sands	566
11.16.6	**Wastewater**	567
11.16.6.1	Wastewater from Shale Gas and Shale Oil	567
11.16.6.1.1	Disposal of wastewater from shale gas and shale oil	567
11.16.6.1.2	Treatment and reuse of wastewater from shale gas and shale oil	568
11.16.7	**Synthesis**	568
Acknowledgments		569
References		569

11.16.1 Introduction

A view of the carbon cycle that disregards human activity misses the fastest flow within the cycle – transport of carbon from Earth's crust to the atmosphere through the burning of fossil fuels. Humans now move carbon from the crust to the atmosphere at a greater rate than plants and the oceans absorb carbon dioxide from the atmosphere. In a more inclusive view of the carbon cycle, anthropogenic climate change can be seen as a problem that has emerged from the interruption of an established balance of carbon among a few stable reservoirs.

The role of human activities in other material cycles, such as alteration of freshwater flow among reservoirs of the hydrologic cycle, is only beginning to receive attention similar to that of the carbon cycle. Interest in freshwater resources is driven partly by the ever-increasing water demands of a growing population and by climate change (see, e.g., Vorosmarty et al., 2000). Until 1996, when Postel et al. (1996) calculated that global human appropriation of freshwater could outpace freshwater availability, freshwater resources had been managed at a scale no larger than that of a river basin. As consciousness of the scarcity of global freshwater resources increased, however, discussion emerged over the disproportionate amounts of available freshwater among countries and the flow of freshwater between these countries in the form of virtual water – the water contained in, or used to manufacture, a traded commodity (Allan, 2011; Hoekstra and Chapagain, 2008; Hoekstra and Hung, 2005; Hoff, 2009; Mekonnen and Hoekstra, 2011). As development of unconventional resources increases around the world, cognizance of how consumptive water use for fuel production will affect freshwater flows and availability is paramount, especially in water-scarce regions.

The production of conventional fossil fuels has largely relied on the pressure difference between the fluids in the pores of reservoir rocks at depth and atmospheric pressure at the surface to bring freely flowing hydrocarbons up from the reservoir. Conventional reservoirs, in general, have not required large volumes of freshwater for extraction, but can, over their lifetimes, produce large quantities of brine in addition to hydrocarbons. Conversely, methods for extracting fossil fuels from unconventional reservoirs tend to require much higher levels of freshwater input, which can lead to larger volumes of wastewater returned from the well at the onset of production and, therefore, a higher risk of immediate ground and surface water contamination. Important ethical questions surround the export of water-intensive technologies for producing unconventional reservoirs to water-scarce regions. Ongoing debate about the trade in unconventional resources and technology should be informed by the trade of virtual water contained in these resources. If unconventional resources (e.g., shale gas) will be the basis of a transition to renewable energy sources that may take several decades, then the impacts of unconventional hydrocarbon extraction on freshwater resources must be taken into account. The objective of this chapter is to review some of the unconventional reservoirs likely to be used in the next few decades and examine the impacts of their production on freshwater resources.

A Note on Hydrocarbon Resource Terminology and Estimates

The terminology for fossil fuel resources varies considerably. Generally, the term 'resource' refers to fossil fuels that are likely to exist in Earth's crust, based on geologic evidence. 'Reserves,' however, are fossil fuels that can be

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00202-5

mined economically using today's technology. These are the definitions that will be used in this chapter.

Estimates of resources and reserves made by different organizations are often inconsistent due to differing assumptions and methodologies. This chapter uses a variety of sources for estimates of resources and reserves that are not always directly comparable.

11.16.2 Unconventional Resources

Unconventional resources can have a variety of definitions. For example, unconventional resources may refer to any resource for which production technology is not yet mature, extraction is not yet economical, environmental issues are not yet resolved, or any combination of these (Oristaglio, 2012). Unconventional fossil fuels may also refer to resources extracted directly from source rock, rather than reservoir rock. More generally, unconventional resources and methods of extraction are less well understood than conventional resources since unconventional resources have not been as significantly studied and exploited to date. Two methods of classification of hydrocarbons can be used to contrast unconventional and conventional hydrocarbons: the McKelvey diagram and the natural resource pyramid.

McKelvey, the director of the US Geological Survey from 1971 to 1978, first described the McKelvey diagram in 1972 (McKelvey, 1972). At the time the diagram was proposed, there existed an array of both geologic and economic classifications of fossil fuel resources meant to inform policies on location and timing for extraction, as well as decisions to import additional resources. To help provide a better framework for these decisions, and to identify areas where additional information was needed, McKelvey conceptually combined geologic and economic classifications into one diagram. The original McKelvey diagram shows that the degree of geologic certainty concerning the availability of the resource increases from right to left and the economic feasibility of recovery increases from bottom to top (Figure 1). Within this diagram, conventional resources that are geologically well defined would be located in the left side of the reserve box in the region where resources will be exploited, whereas the bottom half of the diagram, characterized by marginal economic feasibility of recovery, would be composed of unconventional resources. (Some conventional resources that are hard to access or unconfirmed geologically may also make up a smaller portion of the bottom of the diagram.)

Over time, as technology and knowledge improve and the price of fuels changes, resources can migrate within the diagram. For example, the relative increase in natural gas prices at the start of the 20th century caused known shale gas resources to become more attractive. This led to increased exploration and delineation of shale gas reservoirs and further development of technology necessary to make shale gas production economically viable. The McKelvey diagram provides a conceptual framework not only for making distinctions among conventional and unconventional fuels but also for tracing the progress of an unconventional resource to a known reserve and, eventually, a conventional reserve. Nevertheless, this diagram leaves out some critical dimensions along which resources must be evaluated.

The second conceptual method that can be used to compare conventional and unconventional fossil fuels is the resource pyramid. The resource pyramid improves on the McKelvey diagram by including a dimension that represents the relative size of each resource. It is organized such that the top of the pyramid represents the highest quality and easiest to extract resources, while the bottom of the pyramid represents the lowest quality

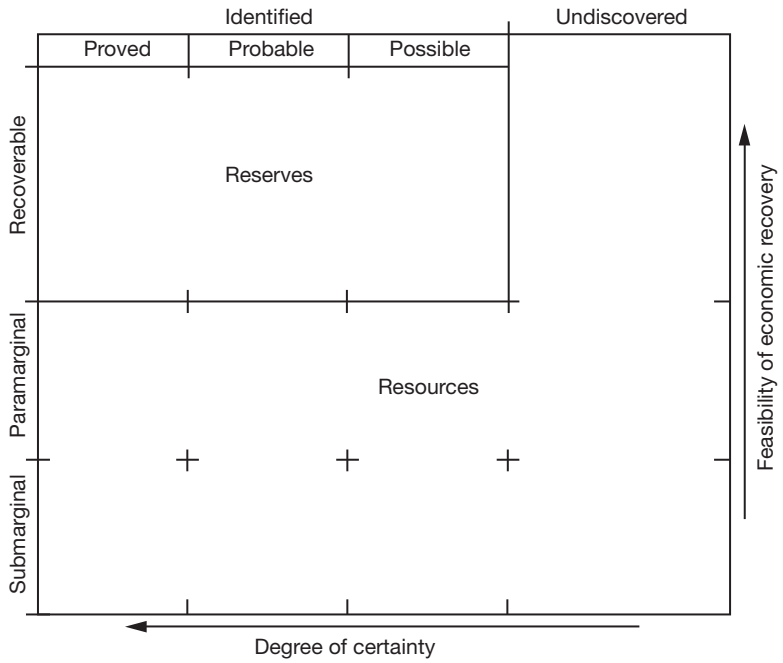


Figure 1 Original McKelvey diagram for fossil fuel resources. Reproduced from McKelvey VE (1972) Mineral resource estimates and public policy. *American Scientist* 60: 32–40.

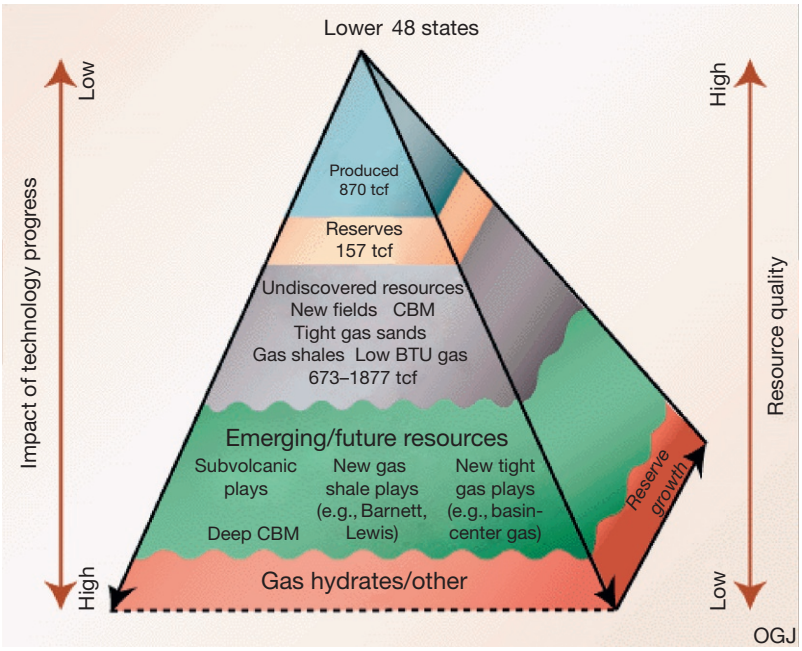


Figure 2 Natural gas resource pyramid. Reproduced from Kuuskraa VA, Schmoker JW, and Dyman TS (1998) Diverse gas plays. *Oil and Gas Journal*. Available online: http://www.ogj.com/articles/print/volume-96/issue-23/in-this-issue/exploration/diverse-gas-plays-lurk-in-gas-resource-pyramid.html.

and most difficult to extract resources. According to this framework, conventional fuels appear at the top of the pyramid and unconventional resources appear in the middle and bottom. For example, in the natural gas resource pyramid, conventional natural gas is located at the top, and unconventional resources, including shale gas plays, are located toward the bottom (**Figure 2**; Kuuskraa et al., 1998). Importantly, the third dimension of the pyramid, identified as 'reserve growth,' demonstrates the relative size of each resource in the pyramid. This is a critical piece of information as it demonstrates that unconventional resources that appear toward the bottom of the pyramid have the potential to provide a much larger supply of fossil fuels than conventional resources.

An examination of the world's total oil reserves provides a good example of a resource pyramid in which the highest-quality reserves are smallest, while lower-quality resources are abundant. The economic value of oil, represented as quality and ease of extraction in the resource pyramid, is generally determined by the density and viscosity of the oil. The denser and more viscous the oil, the lower its economic value because it is more costly to extract and refine. Nevertheless, the world's supply of heavy oils is much larger than that of light oils. Of the world's total estimated oil reserves, only 30–40% are composed of conventional oil (Alboudwarej et al., 2006; IEA, 2013). The majority of the world's oil is composed of heavy oil, extra heavy oil, and oil sands and bitumen (**Figure 3**; Alboudwarej et al., 2006). The world's conventional and unconventional natural gas recoverable resources are closer in size with each estimated to be approximately 400 trillion cubic meters (1.41×10^{16} ft^3) (IEA, 2012).

One important shortcoming of both the McKelvey diagram and the resource pyramid is that neither accounts for the geopolitical availability of these resources. In comparison to conventional fuels that are disproportionately available in the Middle East and Russia, unconventional resources are widely available in the United States and in countries, such as Canada, with which the United States has a strong relationship. In this chapter, we will specifically be discussing unconventional resources in North America. In particular, we will focus on the two groups of unconventional resources for which exploitation is already under way in North America: shale gas and tight oil, and heavy oil and oil sands. For each resource type, a brief description of the resource, size, and method of extraction will be given, followed by a discussion of proven and potential impacts on freshwater resources from extraction.

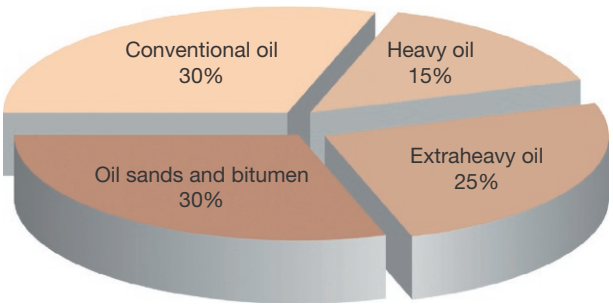


Figure 3 Total world oil reserves. Heavy oil, extraheavy oil, and bitumen, make up about 70% of the world's total oil resources of 9–13 trillion bbl. Reproduced from Alboudwarej H, et al. (2006) Highlighting heavy oil. *Oilfield Review* 18(2): 34–53, Summer 2006. Available online: http://www.slb.com/~/media/Files/resources/oilfield_review/ors06/sum06/heavy_oil.pdf.

11.16.3 Shale Gas and Shale Oil

Shale gas and tight oil (also called light tight oil) are natural gas and low-viscosity crude oil that are trapped within low-permeability formations, such as shale. Gas and oil were produced in the shale formation as a result of the diagenesis of organic matter trapped in fine marine sediments that were compacted into shale formations. Due to the very low permeability – on the order of nanodarcies – of the shale, gas and oil do not flow into a well drilled through the shale formation at economic rates. As a point of comparison, permeability of conventional oil reservoirs is on the order of millidarcies, and permeabilities greater than 100 millidarcies are considered excellent for production (Hyne, 2001). Therefore, to allow for the flow of hydrocarbons in low-permeability formations, the shale formation must be stimulated with a mixture of water, sand, and chemical additives. This process of stimulating the shale formation with highly pressurized fluid is called hydraulic fracturing and has become a controversial practice, in part due to the risks it poses to freshwater resources.

While estimates of total US shale gas and tight oil resources are frequently revised, the current estimates are that US basins hold a total of 13.6 trillion cubic meters (482 trillion cubic feet) of unproved technically recoverable shale gas resources and 33.2 billion barrels of unproved technically recoverable tight oil resources (EIA, 2012c). These US resources are found within shale formations throughout the United States (**Figure 4**; EIA, 2011). Over the past few years, rates of extraction of shale gas and tight oil in the United States have increased dramatically. Between 2007 and 2010, shale gas production increased fivefold from 36.6 to 151.1 billion cubic meters per year (1293–5336 billion cubic feet per year) (EIA, 2012a). As a result, shale gas is projected to comprise 50% of all domestic natural gas production by 2040 (EIA, 2013). Similar increases in US tight oil production have also occurred. Tight oil production more than tripled between the outset of 2009, when the production rate was 250 000 barrels per day, and the end of 2011, when production rates were at 900 000 barrels per day. Remarkably, over 80% of the tight oil production at the end of 2011 was from two tight oil plays, the Bakken and Eagle Ford shales (EIA, 2012b).

11.16.3.1 Shale Gas and Shale Oil Extraction

Hydraulic fracturing is only one step in the process of shale gas and tight oil extraction. The process begins with site selection

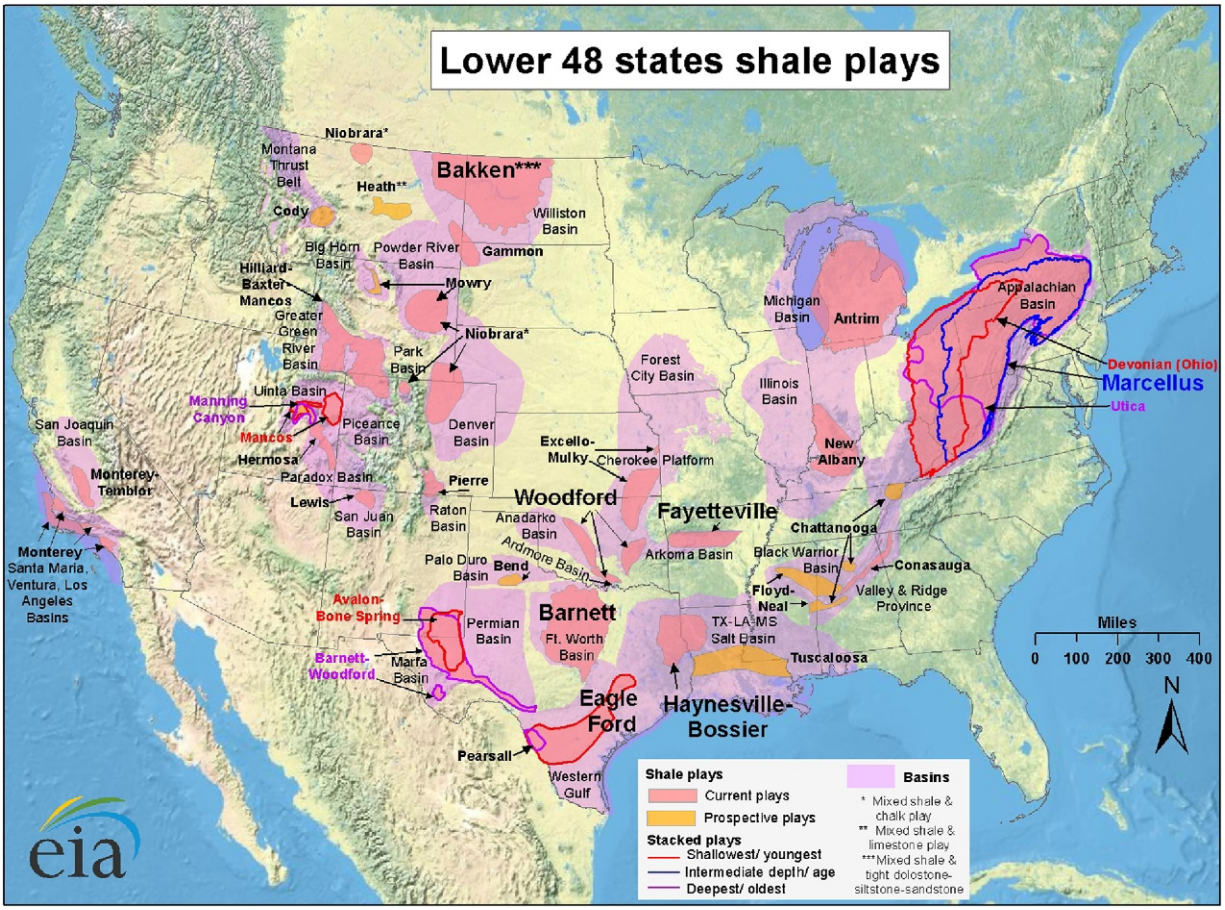


Figure 4 Shale plays in the conterminous United States. Reproduced from EIA (2011) *Natural Gas Maps*: *Exploration*, *Resources Reserves and Production*. Available online: http://www.eia.gov/pub/oil_gas/natural_gas/analysis_publications/maps/maps.htm.

and clearing of approximately nine acres to accommodate the well site and associated infrastructure (Johnson, 2010). The vertical portion of the well is then drilled to the depth of the shale formation, through the water table. During the drilling process, a drilling fluid, which may be water, oil, or synthetic-based in combination with clays, weighting agents, and other additives, is used to lubricate the drill. Once the vertical section of the well is drilled, casing strings, steel pipes that line the borehole and protect freshwater aquifers from deeper brine- or hydrocarbon-bearing zones, are cemented in place. The borehole is then turned, by means of gradual adjustment of the drilling direction, and the horizontal section of the well, also called the 'lateral,' is drilled through the shale formation. The toe, or end, of the lateral can extend as far as 2 miles from the vertical section of the well (EPA, 2011b).

Hydraulic fracturing takes place after the lateral has been completed. Water that has been transported to the well site is mixed with a variety of chemicals to create the frac fluid that will be injected into the wellbore. Some examples of chemicals used are hydrochloric acid, which cleans the wellbore and dissolves minerals around the formation, and petroleum distillate, used to reduce friction between the fluid and pipe and minimize pumping costs (Vidic et al., 2013). Biocides, which are typically added to prevent bacteria growth inside the well, are often a focus of concern among groups opposed to hydraulic fracturing. Other examples of chemicals added to fracturing fluid are listed in **Table 1**. Ultimately, the objective of frac fluid injection is to stimulate the shale layer in a way that maximizes fracture complexity and hydraulic conductivity of the shale formation. To achieve a complex fracture, smaller sections of the lateral, called stages, are isolated and stimulated at increasing frac fluid injection rates sequentially from the toe along the lateral (King, 2010). A portion of the injected water is expelled from the well during a period shortly after hydraulic fracturing. This return flow is termed 'flowback.' The well may continue to yield formation water throughout its entire period of hydrocarbon production. This yield is termed 'produced water.' The amount of water that returns to the surface varies by site and shale formation, but any flowback or produced water must be stored at the site until it can be treated or transferred for off-site treatment, disposal, or reinjection in another hydraulic fracturing operation. Once a well is no longer economically productive, it is either refractured or plugged.

11.16.4 Heavy Oil and Oil Sands

Heavy oil and oil sands (also called bitumen or tar sands) are characterized as crude oil resources with high enough specific gravity and viscosity such that they will not flow without some sort of stimulation. In contrast to shale gas and tight oil, which are mined from source rock, heavy oil and oil sands have migrated from their source rock into petroleum traps. Petroleum traps for heavy oil and oil sands, however, need not be as leakproof as those for lighter hydrocarbons as it is often the oil's high density and viscosity that keep the resource in place. Heavy oils are characterized according to American Petroleum Institute (API) gravity (degrees API = [[141.5/specific gravity of oil [using density of oil at 60 °F and water at 60 °F]] − 131.5]), an arbitrary scale that allows for easy comparison of the specific gravity of different oils (EIA, 2014). Generally, heavy oil has an API gravity of 22° or below. Oils with an API gravity of <10° are considered very heavy oils, and tar sands are distinguished by a viscosity several orders of magnitude greater than that of heavy or very heavy oil (ANL, 2008).

As discussed earlier, heavy oil and oil sands and bitumen make up over half of the world's oil resources. US heavy oil resources are approximately 100 billion barrels of oil in place (DOE, 2006). However, the technology necessary to mine heavy oil is generally not economical, except at shallow depths in very large basins. For this reason, oil sands are not as economically attractive in the United States as they are in Canada's vast oil sands reserves. The largest US oil sands deposits are in Alaska, Utah, and Alabama (**Figure 5**; Humphries 2008; IOCC, 1984), while Canada's oil sands deposits are present almost exclusively in three areas of Alberta: Peace River, Athabasca, and Cold Lake. Altogether, these three oil sands regions produce about 1.1 million barrels per day and are projected to produce 3 million barrels per day by 2015 (NEB, 2006).

Table 1 Examples of chemicals used in frac fluid

Additive type	*Example compounds*	*Purpose*
Acid	Hydrochloric acid	Clean out the wellbore, dissolve minerals, and initiate crack in rock
Friction reducer	Polyacrylamide, petroleum distillate	Minimize friction between the fluid and the pipe
Corrosion inhibitor	Isopropanol, acetaldehyde	Prevent corrosion of pipe by diluted acid
Iron control	Citric acid, thioglycolic acid	Prevent precipitation of metal oxides
Biocide	Glutaraldehyde, 2,2-dibromo-3-nitrilopropionamide (DBNPA)	Control bacteria
Gelling agent	Guar/xanthan gum or hydroxyethyl cellulose	Thicken water to suspend the sand
Cross-linker	Borate salts	Maximize fluid viscosity at high temperatures
Breaker	Ammonium persulfate, magnesium peroxide	Promote breakdown of gel polymers
Oxygen scavenger	Ammonium bisulfite	Remove oxygen from fluid to reduce pipe corrosion
pH adjustment	Potassium or sodium hydroxide or carbonate	Maintain effectiveness of other compounds (such as cross-linker)
Proppant	Silica quartz sand	Keep fractures open
Scale inhibitor	Ethylene glycol	Reduce deposition on pipes
Surfactant	Ethanol, isopropyl alcohol, 2-butoxyethanol	Decrease surface tension to allow water recovery

Source: Vidic RD, Brantey SL, Vandenbossche JM, Yoxtheimer D, and Abad JD (2013) Impact of shale gas development on regional water quality. *Science* 340: 826–835.

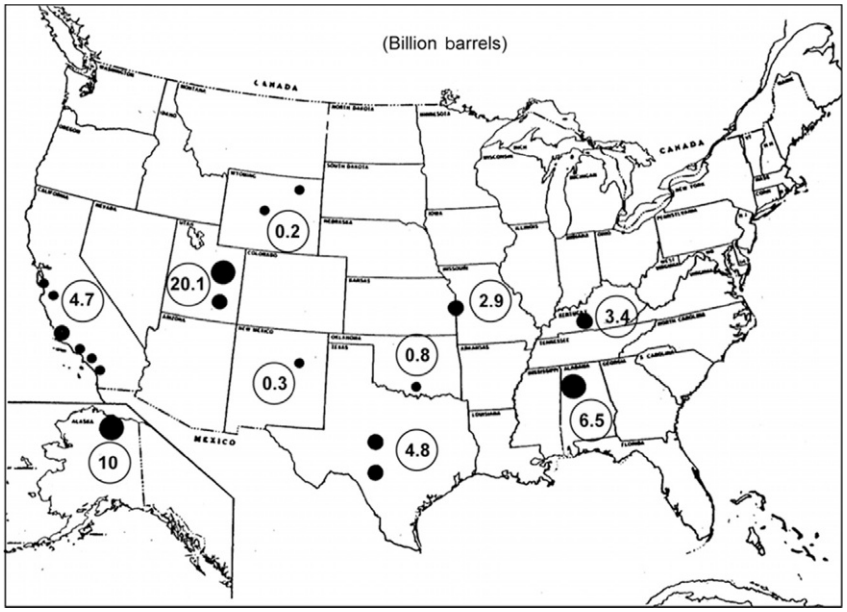


Figure 5 Oil sands deposits in the United States. Adapted from Humphries M (2008) North American oil sands: History of development, prospects for the future. *CRS Report for Congress*. Order Code RL34258. Updated 17 January 2008; IOCC (Interstate Oil Compact Commission) (1984) *Major Tar Sand and Heavy Oil Deposits of the United States*, p. 2. Oklahoma City, OK: IOCC.

11.16.4.1 Heavy Oil and Oil Sands Extraction

Heavy oil and bitumen are recovered either through surface mining, if the hydrocarbon-bearing formation is within 75 m (246 ft) of the surface, or through in situ mining at depth (Kim et al., 2013). The primary in situ stimulation methods used to promote production by lowering the viscosity of heavy oil and bitumen are cyclic steam stimulation (CSS) and a relatively new technology called steam-assisted gravity drainage (SAGD). A few other technologies are used in oil sands production, but this chapter will cover CSS and SAGD as they are among the most widely used and involve water resources.

The process of CSS begins with the drilling of a well to relatively shallow depths (~61 m or ~200 ft) and injecting steam into the well for a period of several weeks. Steam and condensed water heat and soak into the heavy oil or bitumen functioning to decrease its viscosity. Once the viscosity of the oil or bitumen is believed to be sufficiently decreased, the flow direction of the well is reversed to allow for oil to be recovered from the formation. Only 15–30% of resources in place can be extracted after 12–18 cycles of steam injection (ANL, 2008). This limited recovery makes production from even large formations with 100% saturation of heavy oils or bitumen less economically attractive than conventional oil reservoirs. Technological advancements, however, have improved the ability of the oil and gas industry to access these resources. Some of these advancements include marked improvements in methods for measuring the thermal properties of rocks (Chekhonin et al., 2012) and the use of thermal analytic models to improve heat management (Greaser and Ortiz, 2001).

SAGD is characterized by a two-well system. Two wells are drilled into the hydrocarbon-bearing formation parallel to the surface and each other. Steam is injected into the higher-elevation well to decrease the viscosity of the heavy oil or bitumen, which then flows into the lower-elevation well and is pumped to the surface. SAGD is a newer technology than CSS and has some distinct advantages. Whereas CSS recovers as much as 30% of the resource, SAGD can recover up to about 70%. In addition, SAGD requires less steam than CSS, and hence, less water and natural gas are consumed to generate steam (Humphries, 2008).

11.16.5 Water Issues Associated with Unconventional Resource Extraction

The energy–water nexus, a developing area of interest for policymakers, looks at the water resources used in the production of energy and the trade-offs associated with specific energy sources and their respective water and greenhouse gas footprints. It is critical to examine these trade-offs because as the World Policy Institute (WPI) points out, "As energy producers and consumers seek to reduce carbon emissions, water consumption frequently rises because many cleaner forms of conventional and alterative energy are potentially more water intensive" (WPI, 2011). WPI compared the water necessary to both extract and refine fuels that could be used to power an automobile driving from New York City to Washington, DC (**Figure 6**). Surface mining of oil sands clearly requires the largest water resource input. Shale gas production requires only slightly more water than traditional oil, but the majority of water used for shale gas is used during the extraction stage, rather than the fuel refinement stage. This comparison is helpful for conceptualizing differences among fuels, but additional risks to water resources should also be considered in addition to total water appropriation. This section reviews the current knowledge of water resource risks associated with both shale gas extraction and oil sands extraction.

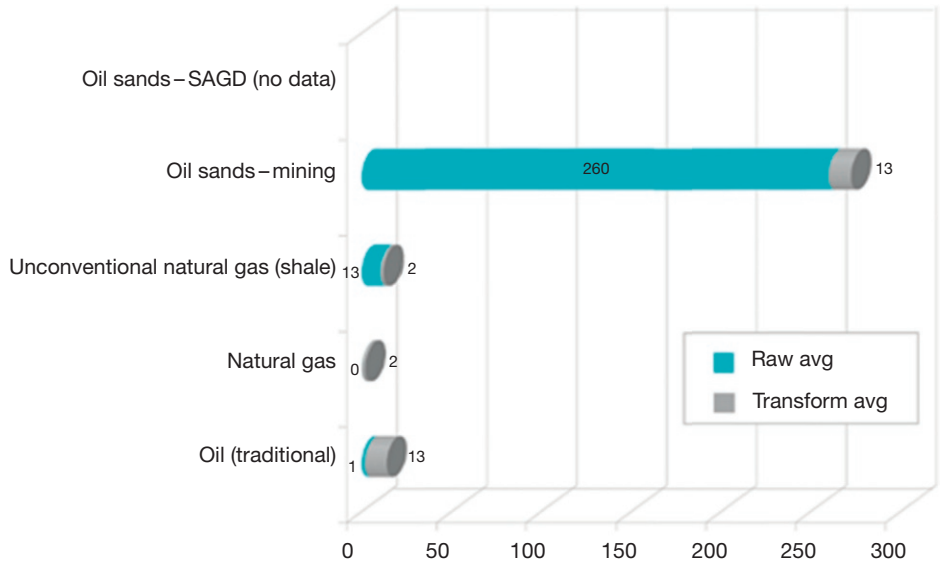


Figure 6 Average (avg) number of gallons of water consumed to produce 1 million BTUs (roughly the energy required to drive from New York City to Washington, DC). Reproduced from WPI (World Policy Institute) in partnership with EBG Capital (2011) The water-energy nexus: Adding water to the energy agenda. World Policy Papers. Available online: http://www.worldpolicy.org/policy-paper/2011/03/18/water-energy-nexus.

11.16.5.1 Water Resources and Shale Gas/Shale Oil Extraction

There are several mechanisms by which shale gas/shale oil extraction could contaminate freshwater or decrease freshwater supply. A growing body of peer-reviewed literature has begun to examine risks to both the quality and the quantity of freshwater. The nascent development of this body of literature and the controversial nature of the topic of hydraulic fracturing have led to much disagreement in the form of official comments on research papers. One of the primary challenges facing researchers in this area is the lack of high-quality, empirical data concerning water use and potential contamination.

11.16.5.1.1 Risk of water contamination with frac fluids and brines

Rozell and Reaven (2011) constructed a statistical model (probability bounds analysis) designed for situations in which data are sparse and parameters are uncertain. This model was used to evaluate the relative risk of freshwater contamination from each step in the process of shale gas extraction from the Marcellus Shale in the Appalachian Basin. Pathways of contamination evaluated by the model included transportation spills, well-casing leaks, leaks through fractured rock, drilling site discharge (or spills), and wastewater disposal. Despite high levels of uncertainty in model results, the model predicted that wastewater disposal was by far the most likely step in the process of shale gas extraction to contaminate freshwater supplies. This result was due, in large part, to the severity of contamination events that could take place should a pond storing produced water fail.

Oil and gas industry representatives assert that there has not been a single case of freshwater contamination from the hydraulic fracturing process itself. While this is true, the inadequate management of produced water has led to documented contamination events. In particular, EPA launched an investigation into groundwater contamination in Pavillion, Wyoming, in response to complaints from homeowners in the region concerning the taste of water from their private drinking water wells (EPA, 2011a). Following a series of sampling events over 2 years, EPA concluded, "Detection of high concentrations of benzenes, xylenes, gasoline range organics, and total purgeable hydrocarbons in groundwater samples . . . indicates that pits are a source of shallow groundwater contamination." While this EPA study came under harsh criticism by industry, it did bring to light that improperly lined pits, or storage ponds, storing produced water are a significant source of risk for water contamination.

A collection of other peer-reviewed studies have tried to address the question of whether formation brines and frac fluids have the potential to migrate from the depth of the gas-bearing shale formation to the overlying drinking water aquifers. Warner et al. (2012) analyzed chemical data and isotopic ratios from shallow groundwater samples collected in the Appalachian Basin and compared them with historical analyses of deep brine samples to determine whether connection pathways exist between deep brines in or near shale formations and shallow groundwater. They found that in northeastern Pennsylvania, the region from which the shallow groundwater samples were taken, there was evidence of conductive pathways. The primary support for this conclusion was that a subset of shallow groundwater samples had Na/Cl, Sr/Cl, Ba/Cl, Li/Cl, and Br/Cl ratios similar to those of the deep brines in the Appalachian Basin. A response to this study, however, suggested that Marcellus Shale brine is unlikely to migrate toward the surface because what little brine is present in the Marcellus Shale, which has a water saturation of $23 \pm 10\%$, is held in place by capillary tension (Engelder, 2012).

Myers (2012) developed a model to simulate the hydrodynamic regime that might lead to communication between deep brines within or near shale formations and shallow groundwater aquifers. Myers (2012) used a numerical model to simulate

the process of hydraulic fracturing and quantify the changes in the hydraulic gradient that would result both at depth and in the shale and sandstone formations overlying the Marcellus Shale. Myers concluded that groundwater-flow perturbations caused by hydraulic fracturing could, under some conditions, contaminate shallow drinking water aquifers with brines forced upward from deep formations. This conclusion, however, has been widely criticized. In particular, Saiers and Barth (2012) and Cohen et al. (2013) asserted that the model excluded key processes, model boundary conditions were incorrectly prescribed, and the hydrostratigraphy was oversimplified, and thus, Myers' model predictions are unrealistic. It is clear from this controversy that additional study is needed to improve the understanding of the fate of injected frac fluid and potentially displaced formation brines that do not return to the surface (Vidic et al., 2013).

11.16.5.1.2 Risk of water contamination with methane

Lack of scientific consensus also exists in the ongoing debate over whether stray gases resulting from shale gas production have the potential to contaminate drinking water wells with methane, the primary component of natural gas. While methane in homeowners' wells is not poisonous, high concentrations of methane in water wells can pose a combustion risk. The US Department of the Interior classifies the action level threshold for methane in drinking water wells to be 10 mg methane per liter of water. There are several ways in which shallow drinking water aquifers could potentially be contaminated by methane, one of which has nothing to do with hydraulic fracturing. In some cases, methane already present in the shallow subsurface close to drinking water wells could migrate a short distance to drinking water wells. This type of methane present in the shallow subsurface is sometimes attributed to bacterial activity in the soil and is therefore called biogenic methane. Contamination of drinking water wells with thermogenic methane, the methane that is present in shale formations, may happen in one of three general ways. First, thermogenic methane from deeper source rock could migrate upward over geologic time to shallow drinking water aquifers, potentially through naturally occurring fracture networks. The other two pathways of contamination require some additional discussion of well casings.

A well casing is a steel pipe that is used to line a well drilled for gas production. Close to the surface, casings are nested such that there are several layers of steel pipe and cement between the gas flowing to the surface and shallow drinking water aquifers (**Figure 7**; API, 2009). Each of these casings, or steel pipes, is cemented in place by pumping cement down the interior of the pipe and then continuing to pump cement down the well until the cement wraps around the bottom of the casing and appears at the surface on the exterior of the pipe. Though the API publishes industry standards for oil and gas casing, each state is responsible for regulating the details of well casings including length of each casing, type of cement used, and quality control procedures after the casing is completed (FracFocus, 2013).

The second pathway that could allow thermogenic methane to migrate to shallow aquifers is through the annulus – the annular space that exists between the outermost cement layer and the formations surrounding the casing. Finally, it has been

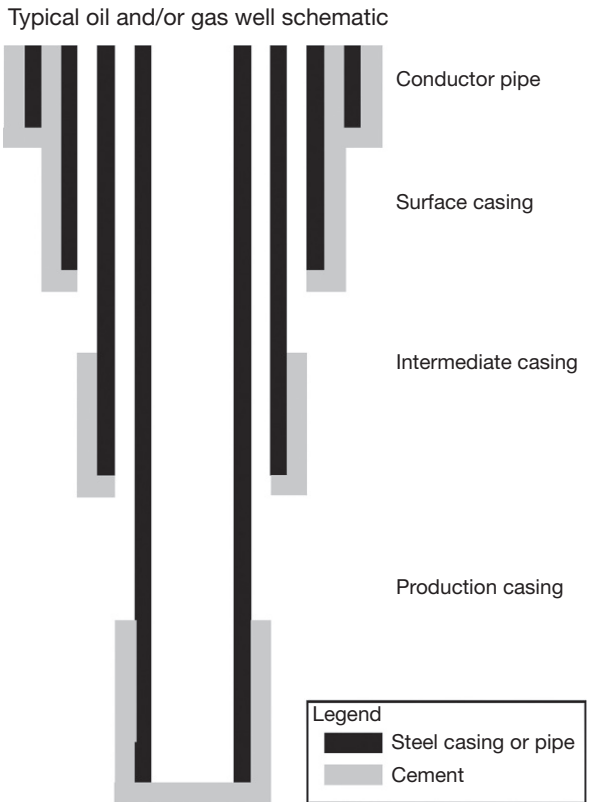


Figure 7 Example of well casing for a shale gas well. Reproduced from API (American Petroleum Institute) (2009) *Hydraulic Fracturing Operations—Well Construction and Integrity Guidelines*, 1st edn. Washington, DC: API. API Guidance Document HF1, October 2009. Available Online: http://www.shalegas.energy.gov/resources/HF1.pdf.

suggested that the pressure on the well casing during stimulation could lead to cracks in the cement. These cracks are the third pathway through which methane could migrate toward the surface.

Osborne et al. (2011) were among the first to analyze water samples from drinking water wells in active drilling areas (within 1 km of at least one unconventional gas well) and nearby nonactive areas of the Marcellus Shale underlying northeastern Pennsylvania and New York. The study found that average and maximum methane concentrations in drinking water wells increased as proximity to active gas production sites decreased. One major critique of this study was that water samples were not taken both before and after drilling began and the comparison of methane concentrations between different wells was insufficient to show that shale gas was the cause of methane in drinking water. Furthermore, other studies have documented the occurrence of methane in drinking water wells prior to hydraulic fracturing. Boyer et al. (2012) compared pre- and postdrilling methane concentrations in 48 water wells in Pennsylvania and found no increase in methane concentrations following hydraulic fracturing (except for one water well) and no correlation between methane concentration and distance to the nearest Marcellus Shale site.

Other critiques of Osborne et al. (2011) pointed out that factors not controlled for in the experimental design would

impact methane concentration in water wells. In particular, methane concentrations are thought to be affected by topographic location (whether the water well is in a valley) and distance from the Appalachian structural front (ASF), a proxy for tectonic deformation and associated thermal maturity. Molofsky et al. (2013), for example, took 1701 samples from water wells in northeastern Pennsylvania and found that methane concentrations are higher in wells located in valleys than in wells in upland areas. In response to suggestion of alternative controls on methane concentration, Jackson et al. (2013) collected samples from 141 drinking water wells in northeastern Pennsylvania and recorded the topographic location and distance from ASF for each sample. Multiple linear regression analysis was then used to determine the significance of topographic location, distance from ASF, and distance from active shale gas development on methane concentrations. Jackson et al. (2013) found that proximity to active shale gas production sites and proximity to ASF were significant predictors of methane concentration in drinking water wells.

Further confounding the question of whether hydraulic fracturing can contaminate drinking water wells with methane is the variation in geology among shale plays. Warner et al. (2013) in conjunction with USGS analyzed groundwater samples from drinking water wells above the Fayetteville Shale in Arkansas and found no statistical difference between water from wells within 1 km of gas wells and water from wells >1 km from gas wells. In the seven wells where methane concentrations were >0.5 mg l^{-1}, a carbon isotope analysis revealed that the methane was likely of biogenic origin (Kresse et al., 2011).

11.16.5.1.3 Water resource appropriation and shale gas/shale oil extraction

Water demand for hydraulic fracturing varies with shale formation, well depth and length, frac fluid properties, and fracture job design (EPA, 2011b). Four of the major shale plays, Barnett, Fayetteville, Haynesville, and Marcellus, require on average 2.3, 2.9, 2.7, and 3.8 million gallons of water per well, respectively (EPA, 2011b). Direct comparison of total water appropriation in different regions, such as a comparison between water use in Pennsylvania and that in Texas, is quite difficult because of the differences in water use per well, number of wells stimulated, percentage of injected water that returns during well production, and rates of recycling. Where freshwater resources are abundant, freshwater is generally less expensive for companies to use for fracturing instead of water with higher salinity. With thousands of wells drilled each year, and the rate of drilling new wells expected to increase, shale gas and shale oil extraction has the potential to consume a significant amount of freshwater throughout the country.

Studies that have quantified freshwater appropriation for shale gas extraction have compared total freshwater withdrawals for hydraulic fracturing to withdrawals for other uses or calculated water efficiency for shale gas production. Overall, water use for fossil fuel production, conventional or unconventional, is significantly less than other water-intensive industries such as agriculture. In Texas, where there are several major plays producing shale gas, the volume of water used for all types of mining in 2008 was among the smallest of all industries (**Figure 8**; Nicot and Scanlon 2012b). Within the mining industry, however, hydraulic fracturing in Texas is among the largest water users (**Figure 9**; Nicot and Scanlon, 2012b). Nicot and Scanlon (2012a) quantified total water use for the three primary shale plays in Texas and found that withdrawals for hydraulic fracturing made up less than one percent of all water withdrawals in the state (Nicot and Scanlon, 2012a). But, this study and others examining total water use for shale gas extraction have not analyzed impacts on freshwater resources resulting from the location and timing of these withdrawals.

Goodwin et al. (2012) examined the water use and intensity of oil and gas extraction in the Wattenberg field in northeastern Colorado. The study looked at water use for 445 wells – 386 vertical and 59 horizontal – operated by Noble Energy. They found that vertical wells use an average of 380 000 gallons of water and horizontal wells use an average of 2.8 million

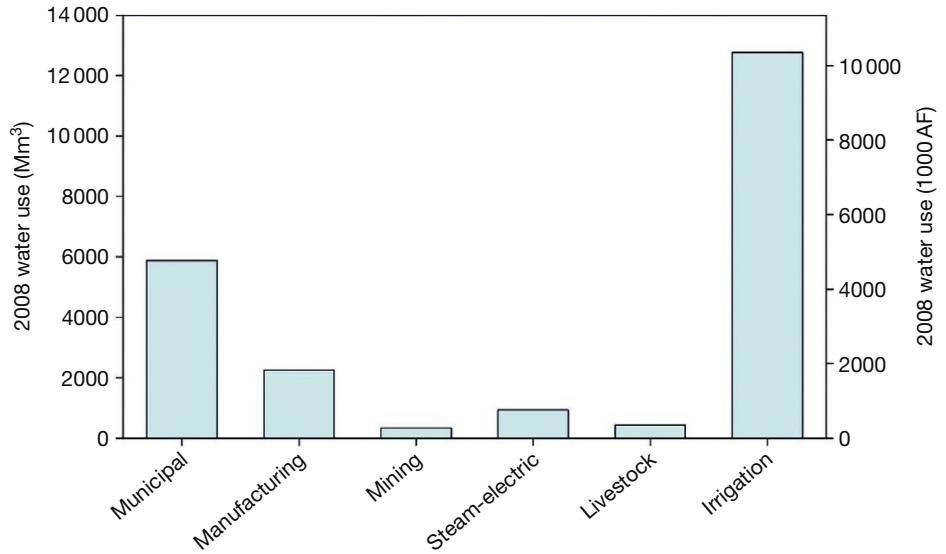


Figure 8 2008 Texas water use by industry. Reproduced from Nicot JP and Scanlon BR (2012) Water use for shale-gas production in Texas, US [supporting information]. *Environmental Science and Technology* 46: 3580–3586.

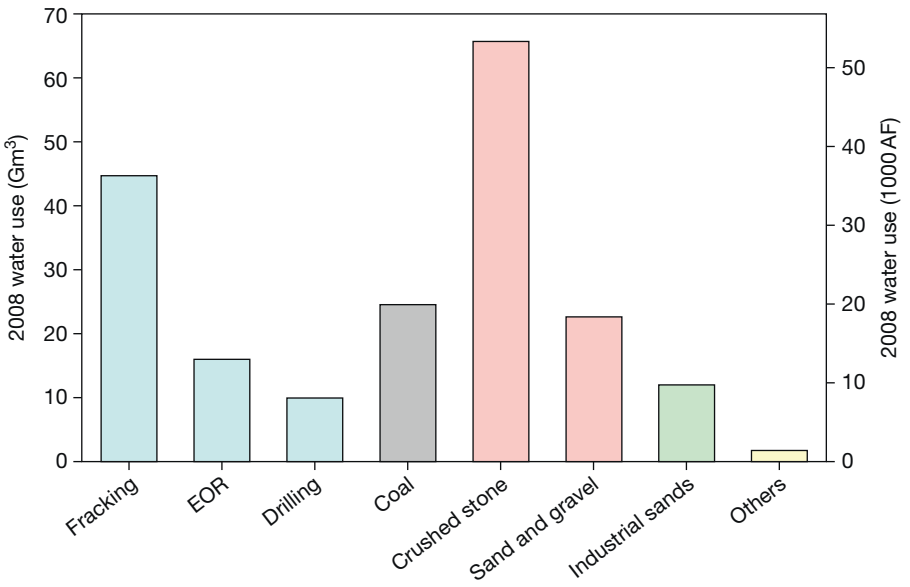


Figure 9 2008 Texas water use by mining category. Reproduced from Nicot JP and Scanlon BR (2012) Water use for shale-gas production in Texas, US [supporting information]. *Environmental Science and Technology* 46: 3580–3586.

gallons of water. Despite higher overall water use for the completion of horizontal wells, water intensity – defined as volume of water used per unit energy provided by the natural gas recovered – was higher for vertical wells than horizontal wells. In other words, fracturing of horizontal wells is ultimately a more water-efficient method for extracting shale gas. On average, water intensity was 6.9 gallons per million British thermal units (MMbtu) for vertical wells and 4.3 gallons per MMbtu for horizontal wells (Noble Energy and Colorado State University, 2012). Goodwin et al. (2012) also reported water intensity associated with shale gas extraction was lower than that for coal, most types of oil-bearing reservoirs, and biofuels.

A more complete picture of water appropriation for energy production can be obtained by tracing water consumption through each phase of energy generation associated with a particular fuel source. Grubert et al. (2012) compared water intensity for three different fuels used in Texas for energy generation. Water intensity was analyzed for fuel extraction, power plant cooling, and power plant emissions control phases of each fuel's lifecycle. Based on the fuels included in the study, Texas lignite coal, conventional natural gas, and shale gas, it was found that switching from traditional coal-fired power plants to natural gas combined cycle plants could reduce freshwater consumption in Texas by about 53 billion gallons per year (Grubert et al., 2012). While understanding the water footprint of each fuel through energy production is imperative to decisions concerning fuel switching, even this broad view of water consumption for electricity production does not take into account water use for treatment of waste products that result from fuel extraction.

This review of the current literature concerning freshwater risks from shale gas extraction demonstrates that the understanding of hydrologic impacts of hydraulic fracturing is still developing. Areas of greatest concern and ongoing research efforts include proper handling and treatment of produced water, methane contamination of drinking water, and total consumptive water use in water-scarce regions.

11.16.5.2 Water Resources and Heavy Oil/Oil Sands

One focus of technological innovation for enhanced recovery of heavy oil and bitumen is finding a way to produce steam more efficiently. Steam production for thermal enhancement of heavy oil reservoirs is the largest cost in any thermal operation (Greaser and Ortiz, 2001) and represents a large sink of local water resources. Producers in California are estimated to use 560 million barrels of water per year for steam injection (ANL, 2008). In return, during production, enormous amounts of produced water return to the surface with the oil. Chevron estimates that for every one barrel of oil produced in the Kern River field, nine barrels of produced water return to the surface. This translates to a total of 900 000 barrels per day of produced water (ANL, 2008). Similarly to shale gas and tight oil, sourcing sufficient water resources for steam production and appropriate disposal of wastewater are two of the primary water issues associated with heavy oil production.

Oil sands development in Alberta, Canada, has the potential to have a significant impact on rivers in the region (Radke, 2006; Schindler et al., 2007). Water use projections based on planned mining operations in Alberta's oil sands suggest that the Athabasca River may not have sufficient flows to meet all the industry demands and provide environmental flows (Radke, 2006). A review of water security issues related to oil sands reported that for every barrel of bitumen recovered by surface mining, two to three barrels of water are withdrawn from the Athabasca River and ~85% of water used in surface mining is recycled (Kim et al., 2013). In situ extraction of bitumen is more efficient and only requires 0.9 barrels of water per barrel of bitumen produced, and up to 95% of water used can be recycled (Kim et al., 2013). Despite relatively high recycling rates, water

security for continued oil sands development is a concern as the production of bitumen is expected to grow. Improving management and monitoring of water withdrawals and produced water recycling and disposal is a challenge that must be resolved to protect both water resources and the continued health of the industry (Kim et al., 2013).

In order to mitigate the water use and potential water contamination issues associated with oil sands, several researchers have called for a more robust monitoring system and further research. In particular, David Schindler, an ecologist at the University of Alberta, authored a comment in Nature to call attention to the lack of responsibility taken to safeguard the health of the Athabasca River and its tributaries (Schindler, 2010). He explains that inadequacies related to scientific leadership of organizations responsible for monitoring river health have led to inconsistent sampling and methodology. Ultimately, the dearth of reliable data has made it difficult to assess the water quality and quantity impacts of large open pits, disruption of surface and groundwater hydrology, and accumulation of liquid waste in tailing ponds associated with tar sands development (Elshorbogy et al., 2005; Kim et al., 2013; Schindler, 2010).

11.16.6 Wastewater

Produced water, whether it comes from a shale gas well or a heavy oil/bitumen well, contains not only chemical additives used to enhance gas and oil production but also naturally occurring and potentially harmful substances absorbed during interaction with shale, oil, and reservoir rock formations. In particular, heavy metals, radionuclides, and very high levels of total dissolved solids (TDS) present in produced water can be harmful to human health and the environment when produced water is not handled correctly. Volumes of produced water vary by type of fossil fuel being produced, method of production, and specific formation. Conventional oil reservoirs, for example, typically produce more water than conventional gas reservoirs because of the higher compressibility and sorption capacity of gas (DOI, 2011). Further, Lutz et al. (2013) compared volumes of water produced by conventional natural gas and Marcellus Shale gas wells and found that Marcellus Shale gas wells produce less water per unit gas recovered. Produced water can be disposed of via underground injection, treated, and released back into the environment, or it can be treated and reused for oil and gas production. Treatment and reuse of produced water is, in theory, the best option as this could help to reduce the appropriation of other freshwater resources; however, the technology for produced water treatment and reuse has not been widely adopted in many unconventional resource plays, perhaps owing to comparatively high costs.

11.16.6.1 Wastewater from Shale Gas and Shale Oil

Following the stimulation of a low-permeability formation via hydraulic fracturing, water that returns to the surface must be collected and recycled or disposed of safely. The decision of whether to dispose of flowback and produced water or recycle it can be described as a balancing act. Proximity of disposal locations, costs of trucking the wastewater, and regulations concerning the amount of water that is allowed to be stored at the site all factor into the decision of whether a company will decide to dispose of or recycle their wastewater (Horner et al., 2011). If reuse is determined to be the best option, the appropriate treatment option is determined, in part, by the composition of the wastewater.

Maloney and Yoxtheimer (2012) analyzed records for waste production and disposal from shale gas wells in Pennsylvania and found the rates of disposal and recycling were different for flowback water (defined here as fluids returning to the surface 20–30 days following fracturing) and produced water (called brine water in this study). In 2011, 89.8% of reported flowback was reused, whereas only 55% of brine water was reused. Total quantities of flowback and brine water produced in 2011 from the estimated 1037 new wells were 7 878 587 and 9 065 470 barrels, respectively. The difference in reuse rates between flowback and brine water is attributed to increased difficulty in treating produced water due to higher TDS (Maloney and Yoxtheimer, 2012).

11.16.6.1.1 Disposal of wastewater from shale gas and shale oil

Underground injection and treatment and release into the environment are the two disposal methods for wastewater that is not reused. Each of these options poses some risk to environmental and human health. Underground injection wells are the primary method of wastewater disposal for all shale plays except the Marcellus shale (Veil, 2010).

Underground injection has led to concerns about increased seismicity. Increases in seismicity have been observed in oil and gas production and disposal regions since these activities began. In particular, Ellsworth et al. (2012) noticed an increase in the rate of $M > 3$ earthquakes in the midcontinent (defined as 85° to 108° W and 25° to 50° N). In the period from 1970 to 2000, these earthquakes occurred at a rate of 21 ± 7.6 per year, and from 2001 to 2008, at a rate of 29 ± 3.5 per year. In the following 3 years, 2009, 2010, and 2011, 50, 87, and 134 earthquakes occurred, respectively (Ellsworth et al., 2012). The largest earthquake under investigation for potential induction through underground injection was an earthquake of moment magnitude 5.7 in central Oklahoma (Keranen et al., 2013). Earthquakes of similar magnitude (5.7) in developing countries have led to significant damage, but large-scale damage is unlikely in North America and northern Europe, where buildings are likely to withstand these forces (Ellsworth, 2013). Further, the majority of injection wells in the United States are characterized as aseismic (Ellsworth, 2013).

The disposal of produced water through treatment and release into the environment puts human health and ecosystems at risk if treatment is inadequate. Prior to spring of 2011, some municipal sewage treatment plants in Pennsylvania were accepting flowback and produced water. In April and May of 2011, however, treatment plants that could not efficiently handle treatment of wastewater with very high TDS concentrations were discouraged from accepting wastewater from shale gas extraction (Maloney and Yoxtheimer, 2012). By January of 2012, transport of any shale gas extraction wastewater to municipal waste treatment plants ceased (Olmstead et al., 2013). Despite limiting the treatment of flowback and produced water to industrial wastewater treatment plants in the Marcellus Shale region, concerns about the water quality downstream from treatment plants remain. Olmstead et al. (2013) examined the effect of effluent

from the treatment of shale gas extraction wastewater on downstream surface water quality. Specifically, concentrations of chloride (Cl^-) and total suspended solids (TSS) were examined in surface water downstream of waste treatment plants. Results found that the upstream density of facilities accepting waste from shale gas development had a statistically significant correlation with increased Cl^- concentration downstream, but no correlation was found with TSS concentration.

11.16.6.1.2 Treatment and reuse of wastewater from shale gas and shale oil

Many companies extracting shale gas are interested in methods of treating and reusing wastewater for hydraulic fracturing. This is not a simple endeavor because the composition of flowback and produced water is highly variable. Composition depends on the chemistry of the frac fluid, the amount of time the water remains at depth, the extent of mixing between frac fluids and formation waters, and the geology of the formation (**Table 2**; Horner et al., 2011). Even within the same shale formation, the composition of wastewater can vary from well to well and over time within a single well. Shale formations are a composite matrix of carbonate, organic matter, siliciclastics, and salt crystals (Horner et al., 2011), so the greater the fluid residence time, the higher the TDS of the flowback and produced water. High TDS concentrations in combination with frac fluid chemical additives, naturally occurring radioactive material, bacteria, volatile organics, hydrocarbons, and ammonia are all present in flowback and produced water. This diversity of dissolved and suspended constituents makes these oil and gas wastewaters difficult to treat (Horner et al., 2011).

When flowback water and produced water are to be reused for well stimulation, suspended solids and soluble organics are removed and the rest is combined with freshwater until it reaches a point where it can be mixed to form a new frac fluid. Many of the frac fluid additives perform best when mixed with freshwater. There are some alternative chemicals that can be used for frac fluid with higher TDS, but these additives are more costly, and it is therefore generally preferable to reduce the TDS concentration as much as possible prior to reuse (Horner et al., 2011).

Table 2 Average flowback concentrations from three shale plays

	Fayetteville	*Marcellus*	*Barnett*
Na (mg l^{-1})	5362.6	24445.0	12453.0
Mg (mg l^{-1})	77.3	263.1	253.0
Ca (mg l^{-1})	256.3	2921.0	2242.0
Sr (mg l^{-1})	21.0	347.0	357.0
Ba (mg l^{-1})	0.8	679.0	42.0
Mn (mg l^{-1})	0.5	3.9	44.0
Fe (mg l^{-1})	27.6	25.5	33.0
SO_4 (mg l^{-1})	149.4	9.1	60.0
HCO_3 (mg l^{-1})	1281.4	261.4	289.0
Cl (mg l^{-1})	8042.3	43578.4	23797.5
TDS (mg l^{-1})	15219	72533	39570
S.G.	1.010	1.050	1.030

Source: Horner P, Halldorson B, and Slutz J (2011) *Shale Gas Water Treatment Value Chain – A Review of Technologies, Including Case Studies*. SPE 147264.

Current wastewater treatment technologies can be grouped into three broad categories: membrane filtration, thermal treatment, and other methods that employ some combination of chemical, biological, and physical filtration (Igunnu and Chen, 2012). Within each of these categories, there are several types of evolving technologies, each with their own advantages and disadvantages. The treatment method chosen largely depends on the end-use goals of the treated effluent, such as whether the effluent will be disposed of into the environment or reused in shale gas production. For effluent that is intended for reuse, a current trend in the industry is to maintain mobile treatment centers that can be moved among well pads (Horner et al., 2011).

One method of membrane filtration, referred to as ultrafiltration, consists of a ceramic filter with pore sizes between 0.01 and 0.1 μm (Igunnu and Chen, 2012). This filtration method is particularly effective at removing TSS, but cannot effectively remove salts. In addition, the rates at which this type of treatment can be used may not be fast enough to keep up with the supply of produced water as ceramic membranes can foul and need replacement. Further, waste products that are removed by the membrane must be removed and treated elsewhere (Igunnu and Chen, 2012).

Thermal technologies employ the heating of water to separate clean water from contaminants and are a good option in regions where energy is inexpensive. Thermal technologies for produced water treatment generally predate membrane filtration technologies, which are more commonly used. But, some new thermal technology is useful in treating highly contaminated water (Igunnu and Chen, 2012). One example of thermal treatment is multieffect distillation that converts the wastewater to steam that can then be condensed and recovered as pure water. The high water quality from the product water is an advantage of this technique, but overall product water recovery is relatively low at only 20–35% (Igunnu and Chen, 2012).

Industrial water treatment plants that accept produced water from shale gas operations tend to use some combination of multiple technologies to target the different types of contaminants (Veil, 2010). As regulations concerning produced water evolve, it remains to be seen which treatment technology or combination of technologies may prove to be most economical for both treatment and disposal and treatment and reuse.

11.16.7 Synthesis

This chapter has explored the relationship between unconventional hydrocarbon development and water resources. The research on this topic is scarce and much that is available remains controversial. Owing to this controversy, it is difficult to draw definitive conclusions regarding the impacts of unconventional oil and gas development on freshwater supply and quality. Nevertheless, some cautious inferences can be made. For example, peer-reviewed research suggests that shale gas extraction within northeastern Pennsylvania has, in a limited number of cases, increased methane concentrations in shallow aquifers. This methane occurrence defies consistent explanation, however. That is, recent studies have also revealed that other anthropogenic activities and natural processes introduce thermogenic methane to freshwater aquifers and, moreover, the

frequency of methane occurrence varies with local geology and between shale plays.

While new science emerges on the sensitivity of freshwater resources to unconventional hydrocarbon extraction, policies for regulating the attendant processes are being proposed and considered (see, e.g., NYDEC, 2011). Federal and state agencies, as well as numerous nongovernmental organizations, are seeking to develop policies and best practices that minimize the risk to the environment and human health. In our view, these policies cannot be fully informed on the basis of current knowledge. Consequently, the dangers of either overregulating or underregulating extraction of shale gas, shale oil, and heavy oil persist. To avoid this problem, much more attention should be devoted to advancing understanding of the freshwater-appropriation patterns and water quality impacts of unconventional fossil fuel development. Establishing baseline groundwater and surface water conditions prior to potential alteration from fossil fuel production, for example, is a key step that should be taken to ensure water resource protection.

Finally, the detrimental climate impacts associated with the burning of fossil fuels are now so widely recognized that the relative greenhouse gas impacts of different fuels are extensively analyzed. If it is the case that a transition to alternative energy will include the use of unconventional fossil fuels, the full picture of environmental impacts, especially freshwater impacts, must be understood to ensure that energy decisions that may result in lasting environmental impacts are well informed. It should be recognized that in a world where climate change is already occurring and the planet is supporting more people than ever before, energy decisions should be informed by a variety of factors beyond those traditionally used.

Acknowledgments

We thank M. Oristaglio for reviewing an early version of this chapter.

References

Alboudwarej H, et al. (2006) Highlighting heavy oil. *Oilfield Review* 18(2): 34–53. Summer 2006. Available online: http://www.slb.com/~/media/Files/resources/oilfield_review/ors06/sum06/heavy_oil.pdf.

Allan T (2011) *Virtual Water: Tackling the Threat to Our Planet's Most Precious Resource*. London: I.B. Tauris & Co.

ANL (Argonne National Laboratory) (2008) Water issues associated with heavy oil production. Available online: http://www.perf.org/images/Archive_HeavyOilReport.pdf.

API (American Petroleum Institute) (2009) *Hydraulic Fracturing Operations—Well Construction and Integrity Guidelines*, 1st edn. Washington, DC: American Petroleum Institute. API Guidance Document HF1, October 2009. Available online: http://www.shalegas.energy.gov/resources/HF1.pdf.

Boyer EW, Swistock BR, Clark J, Madden M, and Rizzo DE (2012) The impact of Marcellus gas drilling on rural drinking water supplies, March 2012. Available online: http://www.rural.palegislature.us/documents/reports/Marcellus_and_drinking_water_2012.pdf.

Chekhonin E, et al. (2012) When rocks get hot: Thermal properties of reservoir rocks. *Oilfield Review*, Autumn 2012. Available online: http://www.slb.com/~/media/Files/resources/oilfield_review/ors12/aut12/2_rocks_hot.pdf.

Cohen H, Parratt T, and Andrews C (2013) Comment on "Potential contaminant pathways from hydraulically fractured shale to aquifers" by T. Myers. *Groundwater* 51(3): 317–319.

DOE (2006) *Fact Sheet*: *U.S. Heavy Oil Resource Potential*. Available online: http://fossil.energy.gov/programs/reserves/npr/Heavy_Oil_Fact_Sheet.pdf.

DOI (2011) *Oil and Gas Produced Water Management and Beneficial Use in the Western United States*. Available online: http://www.usbr.gov/research/AWT/reportpdfs/report157.pdf.

EIA (2011) *Natural Gas Maps*: *Exploration*, *Resources Reserves and Production*. Available online: http://www.eia.gov/pub/oil_gas/natural_gas/analysis_publications/maps/maps.htm.

EIA (2012a) *Natural Gas*: *Shale Gas Production Data*. Available online: http://www.eia.gov/dnav/ng/ng_prod_shalegas_s1_a.htm.

EIA (2012b) *Petroleum and Other Liquids*: *This Week in Petroleum*, 14 March 2012. Available online: http://www.eia.gov/oog/info/twip/twiparch/120314/twipprint.html.

EIA (2012c) *Annual Energy Outlook 2012*. Available online: http://www.eia.gov/forecasts/archive/aeo12/pdf/0383(2012).pdf.

EIA (2013) *Annual Energy Outlook 2013*. Available online: http://www.eia.gov/forecasts/aeo/pdf/0383(2013).pdf.

EIA (2014) *Petroleum & Other Liquids Data*: *Definitions*, *Sources and Explanatory Notes*. Available online: http://www.eia.gov/dnav/pet/tbldefs/pet_pri_wco_tbldef2.asp.

Ellsworth W (2013) Injection-induced earthquakes. *Science* 341: 1225942.

Ellsworth WL, Hickman SH, Lleons AL, McGarr A, Michael AJ, Rubinstein JL, et al. (2012) Are seismicity rate changes in the midcontinent natural or manmade? *Seismological Research Letters* 83(2): 403.

Elshorbagy A, Jutla A, Barbour L, and Kells J (2005) System dynamics approach to assess the sustainability of reclamation of disturbed watersheds. *Canadian Journal of Civil Engineering* 32: 144–158.

Engelder T (2012) Capillary tension and imbibition sequester frack fluid in Marcellus gas shale. *Proceedings of the National Academy of Sciences* 109(52): E3625.

EPA (2011a) *DRAFT: Investigation of Ground Water Contamination Near Pavillion, Wyoming*. Ada, OK: Office of Research and Development, National Risk Management Research Laboratory.

EPA (2011b) *Plan to Study the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources*. Available online: http://water.epa.gov/type/groundwater/uic/class2/hydraulicfracturing/upload/hf_study_plan_110211_final_508.pdf.

FracFocus (2013) *Well Construction and Groundwater Protection*. Available online: http://fracfocus.org/hydraulic-fracturing-how-it-works/casing.

Goodwin S, Carlson K, Douglas C, and Knox K (2012) Life cycle analysis of water use and intensity of oil and gas recovery in Wattenberg field, Colorado. *Oil and Gas Journal*, 10 (5): 48–59. Available online: http://www.ogj.com/articles/print/vol-110/issue-5/exploration-development/life-cycle-analysis-of-water.html

Greaser GR and Ortiz JR (2001) *New Thermal Recovery Technology and Technology Transfer for Successful Heavy Oil Development*. Society of Petroleum Engineers SPE 69731.

Grubert E, Beach F, and Webber M (2012) Can switching fuels save water? A life cycle quantification of freshwater consumption for Texas coal- and natural gas-fired electricity. *Environmental Resource Letters* 7. http://dx.doi.org/10.1088/1748-9326/7/4/045801.

Hoekstra AY and Chapagain AK (2008) *Globalization of Water: Sharing the Planet's Freshwater Resources*. Oxford: Blackwell Publishing.

Hoekstra AY and Hung PQ (2005) Globalisation of water resources: International virtual water flows in relation to crop trade. *Global Environmental Change* 15: 45–56.

Hoff H (2009) Global water resources and their management. *Current Opinion in Environmental Sustainability* 1: 141–147.

Horner P, Halldorson B, and Slutz J (2011) *Shale Gas Water Treatment Value Chain – A Review of Technologies, Including Case Studies*, Society of Petroleum Engineers. SPE 147264.

Humphries M (2008) North American oil sands: History of development, prospects for the future. *CRS Report for Congress*. Order Code RL34258. Updated 17 January 2008.

Hyne N (2001) *Nontechnical Guide to Petroleum Geology, Exploration, Drilling, and Production*, 2nd edn. Tulsa, OK: PennWell Corporation.

IEA (2012) *FAQ's: Natural Gas*. Available online: http://www.iea.org/aboutus/faqs/gas/.

IEA (2013) *Resources to Reserves 2013. Illustration*: *Oil Production Costs for Various Resource Categories*. Available online: https://www.iea.org/w/bookshop/add.aspx?id=447.

Igunnu ET and Chen GZ (2012) Produced water treatment technologies. *International Journal of Low-Carbon Technologies*, 1–21 (open access article), http://dx.doi.org/10.1093/ijlct/cts049.

IOCC (Interstate Oil Compact Commission) (1984) *Major Tar Sand and Heavy Oil Deposits of the United States*. Oklahoma City, OK: Interstate Oil Compact Commission, p. 2.

Jackson R, et al. (2013) Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proceedings of the National Academy of Sciences* 110(28): 11250–11255.

Johnson N (2010) Pennsylvania Energy Impacts Assessment Report 1: Marcellus Shale Natural Gas and Wind. Available online: http://www.nature.org/media/pa/pa_energy_assessment_report.pdf

Keranen KM, et al. (2013) Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence. *Geology* 49(6): 699–702.

Kim YJ, Hipel KW, and Bowman CW (2013) Water security problems in Canada's oil sands. *Canadian Water Resources Journal* 38(1): 61–72.

King GE (2010) *Thirty Years of Shale Fracturing: What Have We Learned?* Society of Petroleum Engineers SPE 133456.

Kresse TM, Warner NR, Hays PD, Down A, Vengosh A, and Jackson RB (2011) Shallow groundwater quality and geochemistry in the Fayetteville shale gas-production area, north central Arkansas. *USGS Scientific Investigations Report* 2012–5273.

Kuuskraa VA, Schmoker JW, and Dyman TS (1998) Diverse gas plays lurk in gas resource pyramid. *Oil and Gas Journal* 96(23): 123–130. Available online: http://www.ogj.com/articles/print/volume-96/issue-23/in-this-issue/exploration/diverse-gas-plays-lurk-in-gas-resource-pyramid.html

Lutz BD, Lewis AN, and Doyle MW (2013) Generation, transport, and disposal of wastewater associated with Marcellus Shale gas development. *Water Resources Research* 49: 647–656.

Maloney K and Yoxtheimer D (2012) Production and disposal of waste materials from gas and oil extraction from the Marcellus Shale play in Pennsylvania. *Environmental Practice* 14: 278–287. http://dx.doi.org/10.1017/S146604661200035X.

McKelvey VE (1972) Mineral resource estimates and public policy. *American Scientist* 60: 32–40.

Mekonnen MM and Hoekstra AY (2011) National water footprint accounts: The green, blue, and grey water footprint of production and consumption. *Value of Water Research Report Series No. 50*. UNESCO-IHE Institute for Water Education.

Molofsky LJ, Connor JA, Wylie AS, Wagner T, and Farhat SK (2013) Evaluation of methane sources in groundwater in northeastern Pennsylvania. *Groundwater* 51(3): 333–349.

Myers T (2012) Potential contaminant pathways from hydraulically fractured shale to aquifers. *Ground Water* 50(6): 872–882.

NEB (National Energy Board) (2006) *Canada's Oil Sands Opportunities and Challenges to 2015: An Update. An Energy Market Assessment*, June 2006. Available online: http://www.neb-one.gc.ca/clf-nsi/rnrgynfmtn/nrgyrprt/lsnd/pprtntsndchllngs20152006/pprtntsndchllngs20152006-eng.pdf.

Nicot JP and Scanlon BR (2012a) Water use for shale-gas production in Texas, US. *Environmental Science and Technology* 46: 3580–3586.

Nicot JP and Scanlon BR (2012b) Water use for shale-gas production in Texas, US [supporting information]. *Environmental Science and Technology* 46: 3580–3586.

Noble Energy and Colorado State University (2012) *Lifecycle Analysis of Water Use and Intensity of Noble Energy Oil and Gas Recovery in the Wattenberg Field of Northern Colorado*. Available online: http://cewc.colostate.edu/wp-content/uploads/2012/04/Water-Intensity-Report-Final.pdf.

NYDEC (New York State Department of Environmental Conservation) (2011) *Supplemental Generic Environmental Impact Statement on the Oil, Gas, and Solution Mining Regulatory Program*, September 2011. Available online: http://www.dec.ny.gov/energy/75370.html.

Olmstead SM, Muehlenbachs LA, Shih JS, Chu Z, and Krupnick AJ (2013) Shale gas development impacts on surface water quality in Pennsylvania. *Proceedings of the National Academy of Sciences* 110(13): 4962–4967.

Oristaglio M (2012) Personal communication with Yale University Geology and Geophysics senior research scientist in New Haven, CT, 14 November 2012.

Osborn SG, Vengosh A, Warner NR, and Jackson RB (2011) Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing. *Proceedings of the National Academy of Sciences* 108(20): 8172–8176. http://dx.doi.org/10.1073/pnas.1100682108.

Postel SL, Daily GC, and Ehrlich PR (1996) Human appropriation of renewable fresh water. *Science* 271: 785–788.

Radke D (2006) Investing in our future: Responding to the rapid growth of oil sands development. *Alberta Energy Oil Sands Secretariat Reports*. Available online: http://www.energy.alberta.ca/pdf/OSSRadkeReportInvesting2006.pdf.

Rozell DJ and Reaven SJ (2011) Water pollution risk associated with natural gas extraction from the Marcellus shale. *Risk Analysis* 32(8): 1382–1393. http://dx.doi.org/10.1111/j.1539-6924.2011.01757.

Saiers JE and Barth E (2012) Comment on "Potential contaminant pathways from hydraulically fractured shale to aquifers" by T. Myers. *Ground Water* 50(6): 826–828.

Schindler D (2010) Tar sands need solid science. *Nature* 468: 499–501.

Schindler DW, Donahue WF, and Thompson JP (2007) *Running Out of Steam? Oil Sands Development and Water Use in the Athabasca River-Watershed: Science and Market based Solutions* [Section 1]. Toronto: University of Alberta. Available online: http://www.ualberta.ca/~ersc/water.pdf.

Veil J (2010) Water management technologies used by Marcellus Shale gas producers. Prepared for US Department of Energy and the National Energy Technology Laboratory. DOE Award No. FWP 49462.

Vidic RD, Brantley SL, Vandenbossche JM, Yoxtheimer D, and Abad JD (2013) Impact of shale gas development on regional water quality. *Science* 340: 826–835.

Vorosmarty CJ, Green P, Salisbury J, and Lammers RB (2000) Global water resources: Vulnerability from climate change and population growth. *Science* 289: 284–288.

Warner N, et al. (2012) Geochemical evidence for possible natural migration of Marcellus formation brine to shallow aquifers in Pennsylvania. *Proceedings of the National Academy of Sciences* 109(30): 11961–11966. http://dx.doi.org/10.1073/pnas.1121181109.

Warner N, et al. (2013) Geochemical and isotopic variations in shallow groundwater in areas of the Fayetteville shale development, North-central Arkansas. *Applied Geochemistry* 35: 207–220.

WPI (World Policy Institute) in partnership with EBG Capital (2011) The water-energy nexus: Adding water to the energy agenda. World Policy Papers. Available online: http://www.worldpolicy.org/policy-paper/2011/03/18/water-energy-nexus.

11.17 Near-Surface Geophysics at the Hanford Nuclear Site, the United States

TC Johnson, Pacific Northwest National Laboratory, Richland, WA, USA
DF Rucker, hydroGEOPHYSICS, Inc., Tucson, AZ, USA
DR Glaser, Washington River Protection Solutions, LLC, Richland, WA, USA

© 2015 Elsevier B.V. All rights reserved.

11.17.1	Introduction	571
11.17.2	Hanford Site Operational History	571
11.17.3	Plutonium Production Process	572
11.17.4	Waste Disposal Operations	572
11.17.5	Seismic Reflection Imaging of Groundwater Flow Boundaries at the Hanford 200 Area	574
11.17.6	Groundwater/River Water Interaction Beneath Hanford 300 Area Infiltration Ponds	575
11.17.7	Vadose Zone Contaminant Mapping at the Hanford 200 East Area B-Complex	580
11.17.8	Long Electrode ERT Imaging of HLW Inside Tank Farms	582
11.17.9	HLW Tank Leak Detection Using Electrical Methods	586
11.17.10	Engineered Vadose Zone Desiccation Monitoring in the Hanford BC Cribs Area	588
11.17.11	Underground Infrastructure Mapping and Site Clearance	590
11.17.12	Long-Term Legacy: The Role of Geophysics	592
References		593

11.17.1 Introduction

From the inception of the Manhattan Project in the early 1940s until the end of the Cold War, the US Government developed and operated a large program to research, construct, test, and maintain a nuclear weapons stockpile. Over the course of the program, 134 distinct geographic sites were developed, occupying a total area of approximately 5.2 million square kilometers (NRC, 2000). These sites and associated facilities hosted numerous functions, ranging from nuclear material mining operations to weapons testing. Many of these operations, particularly in the early years, generated large amounts of both solid and liquid wastes. Policies and practices for managing these wastes resulted in environmental contamination at many of the sites, particularly those involved in refined nuclear fuel production. As the Cold War and nuclear arms race ended, focus shifted from weapons production to environmental cleanup, restoration, and closure in accordance with US environmental laws. The majority of contaminated sites involved in nuclear weapons production are now under administrative control of the US Department of Energy (USDOE) (http://energy.gov/em/cleanup-sites), which is responsible for cleanup activities and, as such, currently manages one of the world's largest environmental cleanup operations.

Of the 107 sites controlled by the USDOE, 90 have been successfully remediated and closed in accordance with US environmental laws. The remaining 17 open sites are those with the most complex and difficult contamination issues to address. The largest of these is the Hanford Site located in south central Washington State, the United States, which was responsible for research and production of weapons grade plutonium. In this chapter, we discuss some of the challenges associated with subsurface contamination and how near-surface geophysical methods are being used to characterize the subsurface and monitor subsurface processes critical for environmental remediation at legacy US nuclear weapons production sites, using the Hanford Site as a specific example. We begin by reviewing Hanford Site history, leading to a discussion of how waste streams generated during the plutonium production years were managed, resulting in legacy subsurface contamination and the need for near-surface geophysics.

11.17.2 Hanford Site Operational History

The Hanford Site located in south central Washington State, the United States, just north of Richland, WA (**Figure 1**), was selected in 1943 as part of the secretive US project to build an atomic weapon (the Manhattan Project) and was thereafter expanded to produce plutonium for the US nuclear weapons arsenal. Hanford's mission required a large industrial complex including nuclear reactor fuel rod manufacturing facilities, nuclear reactors to produce plutonium, chemical processing plants to extract plutonium from irradiated uranium fuel rods, waste management structures and policies to address the large amounts of nuclear and other hazardous wastes generated, and research facilities to support the site mission (DOE/RL, 1999).

The site operated nine plutonium production reactors in the 100 Area aligned along the northern reach of the Columbia River, which provided cooling water for the reactors. The B reactor located in the 100 BC Area was the world's first plutonium production reactor. The 200 Area, located in Hanford's Central Plateau, housed the chemical separation plants for the site, which were used to extract plutonium from irradiated uranium fuel rods. The 200 Area currently houses most of the site's legacy waste storage facilities and the most complex environmental remediation challenges. The 300 Area was responsible for producing uranium fuel rods for the 100 Area reactors and for research operations.

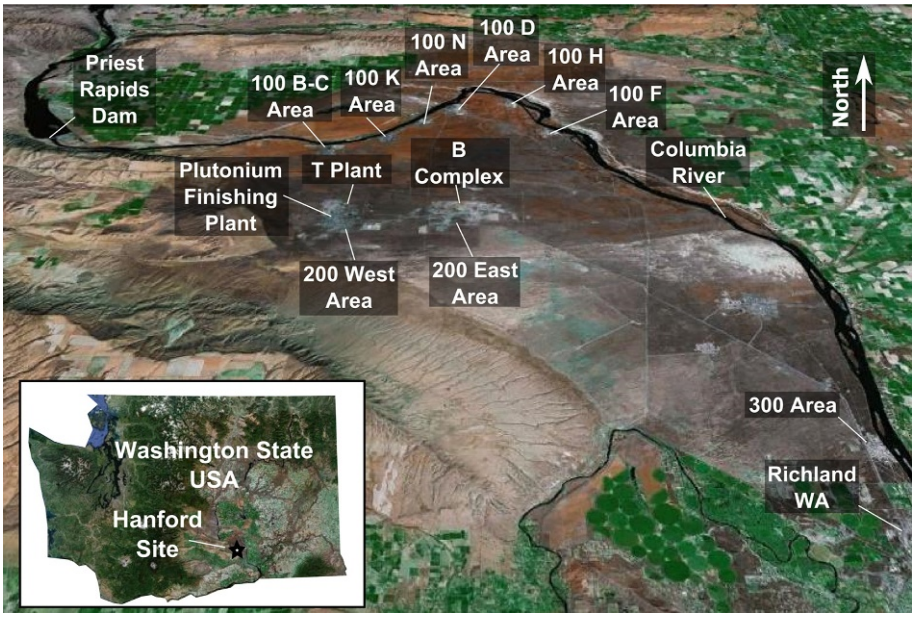


Figure 1 Arial view of Hanford Site in south central Washington, the United States, showing several of the facilities discussed in this document. Nine plutonium production reactors were operated near the Columbia River in the 100 Area. Plutonium separations plants, high-level waste tank farms, and subsurface waste infiltration galleries were operated in the 200 Area. Direct distance from Priest Rapids Dam to Richland, WA, is 61 km. Length of Columbia River from Priest Rapids Dam to Richland, WA, is 105 km.

11.17.3 Plutonium Production Process

Beginning in 1944, Hanford's plutonium production reactors began irradiating uranium fuel rods produced in the 300 Area, ultimately irradiating approximately 100 000 tons of uranium fuel over approximately 50 years (DOE/RL, 1999). The spent fuel was then reprocessed in the 200 Area, where five chemical separations plants were used to dissolve the spent fuel rods and to extract plutonium. Operators dissolved the fuel rods in nitric acid to remove a protective outer jacket and then to dissolve the spent uranium fuel. Plutonium was extracted from the resulting uranium nitrate solution and was sent to the plutonium finishing plant (located in the 200 West Area) to be purified into plutonium metal. Although the plutonium separation and extraction processes were modified over the years to become significantly more efficient with the construction of each new processing plant, each process used complex, toxic, and corrosive chemicals that ultimately produced large amounts of highly radioactive liquid waste (also referred to as high-level waste, HLW). For example, the two original plants (the T and B plants) produced approximately 100 000 gallons of HLW per ton of irradiated fuel, compared to the 250 gallons per ton produced by the final and most efficient processing plant constructed at Hanford, the PUREX plant. The PUREX plant was able to process 33 tons of fuel per day, compared to the 30 tons per month processed by the T and B plants in their most efficient years. Ultimately, the Hanford Site produced nearly 1.04 billion liters of HLW from the five separations plants over approximately 50 years of operation (Gephart, 2003).

11.17.4 Waste Disposal Operations

During late 1940s and 1950s, the Hanford Site was focused on maintaining plutonium production goals. Less focus was placed upon waste management, and the consequences of environmental releases were poorly understood. Recognizing the need to contain high HLW, site operators constructed groups of steel single-shell tanks for HLW storage. Ultimately, 177 of these tanks were constructed subgrade and buried to reduce risk of radiation exposure (**Figure 2**). Operators also used evaporation systems to reduce waste volumes, resulting in a highly concentrated radioactive sludge that was also stored in single-shell tanks. These measures were occasionally inadequate to fully contain the HLW volumes being produced, and consequently, decisions were sometimes made to discharge excess HLW into engineered subsurface infiltration galleries. For example, in the 1940s, HLW generated in the T and B plants was cascaded through a series of three tanks, allowing solid particulates to settle out, before being infiltrated into the vadose zone at the end of the cascade (DOE/RL, 1999). In the 1950s, 18.2 million liters of high-level supernatant waste and 7.2 million liters of evaporator sediment were disposed of in the soil at the Hanford Site.

Due to the corrosive nature of HLW and the temperature variations associated with nuclear fission in HLW streams, many of the single-shell tanks are confirmed or suspected to have failed and leaked (Rodgers, 2013). To mitigate the problem, double-shell tanks were constructed, and wastes are currently being transferred from single-shell to double-shell tanks. Because leaking single-shell tanks presents a continued source

of subsurface contamination, emptying the tanks is currently a primary priority for Hanford Site remediation. The first single-shell tank failure was detected in 1954, only 10 years after construction. Subsequently, an estimated 3.8 million liters of HLW has leaked from single-shell tanks, representing approximately 1 million curies of radioactivity (Gephart, 2003; Rodgers, 2013). Of the 177 single-shell tanks, 67 are known or expected to have leaked, resulting in vadose zone contamination beneath the tanks (Gephart, 2010; Rodgers, 2013).

In addition to HLW streams generated at the chemical separation plants, many Hanford Areas generated copious amounts of radioactive low-level waste (LLW) that was released either into the Columbia River or into engineered subsurface infiltration galleries. Here, LLW is defined as all liquid radioactive wastes that did not meet the prevailing criteria for HLW. The primary source of LLW released into the Columbia River was cooling water used in eight of the nine plutonium production reactors. The final reactor (the N reactor) used a closed circuit of recirculating cooling water. Several methods were also used to infiltrate LLW originating in the chemical separation plants or evaporation facilities into the subsurface at Hanford, but many were discarded due to adverse environmental effects at the surface (Gephart, 2003). Ultimately, the primary method of infiltration used either ponds or shallow buried structures called cribs and trenches. Cribs typically consisted of a series of timbers buried under an earthen mound (Figure 3). Wastes streams were delivered to the timbers through a pipe where they were distributed over a large area to enhance infiltration. Cribs and trenches were used extensively in the 200 Area, with the intent of using the thick vadose zone as a buffer to retain contaminants between the surface and the water table.

Figure 2 Construction of the TX single-shell high-level radioactive waste tank storage facility in the Hanford 200 West Area, 1948. Each tank was ~70 m in diameter and 30 m deep. The vehicle on the road in the upper portion of the figure gives an indication of scale.

Consequent contamination in the deep vadose has reached the water table in some areas and presents a significant risk for future groundwater contamination as discharged effluents continue to migrate downward. Deep vadose zone contamination at the Hanford Site remains one of the most difficult remediation challenges with no permanent cleanup solution identified to date.

Hanford Site operations left a legacy of subsurface contamination that now threatens the groundwater, which ultimately discharges into the Columbia River. Protecting the Columbia River is a primary focus of Hanford Site remediation efforts. Given the nature of subsurface contaminants at Hanford and the risk they pose to human health and the environment, imaging capabilities provided by geophysical methods are increasingly used to inform site remediation operators concerning geology, contaminant distribution, movement, and the natural and engineered processes that govern contaminant behavior. For example, each of the single-shell tank farms has been instrumented with a dense network of wells from which to conduct borehole geophysical logging. Spectral gamma logging within these wells is a primary source of information concerning radioactive contaminant type, distribution, and migration beneath the HLW tank farms. In the following sections, we provide other examples of how near-surface

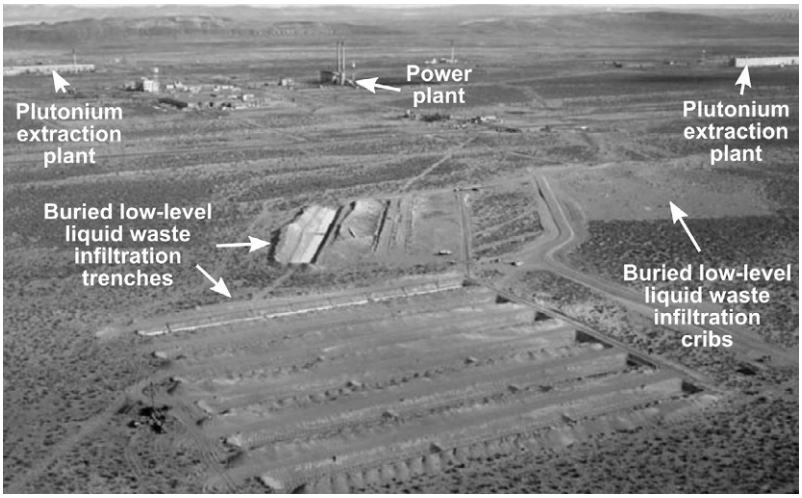

Power plant
Plutonium extraction plant
Plutonium extraction plant
Buried low-level liquid waste infiltration trenches
Buried low-level liquid waste infiltration cribs

Figure 3 Hanford Site BC-Cribs and Trenches (200 East Area) waste infiltration galleries view facing northeast (1955). Cribs and trenches such as these were used to dispose of high volumes of low-level liquid radioactive waste into the vadose zone at several locations in the 200 Area (Figure 1).

geophysical methods are being used to characterize the subsurface and to understand both natural and engineered properties and processes associated with environmental remediation at the Hanford Site.

11.17.5 Seismic Reflection Imaging of Groundwater Flow Boundaries at the Hanford 200 Area

Former waste management practices at the Hanford Site have resulted in numerous groundwater contaminant plumes, some of which have begun to emerge in the last decade as vadose zone contaminants migrate downward and reach the water table (DOE/RL, 2012). Effective management of groundwater contamination requires information concerning the hydrostratigraphic controls governing groundwater movement. Figure 4 shows a generalized stratigraphic cross section through the 200 Area. The unconfined aquifer beneath the central plateau is generally characterized by the high-permeability sands and gravels of the Hanford formation and Ringold formations, underlain by lower permeability basalt and/or lower Ringold mud formations that act as a vertical flow boundary. Of particular importance for this section are the bounding basalt units in the 200 Area, which influence flow paths and in some locations rise above the water table to form a flow barrier. These flow barriers divert groundwater flow and understanding the topography of the basalt contact is critical for accurate prediction of groundwater flow and contaminant transport. Although the location of the basalt boundary has been identified through drilling at many locations, the topographic relief of the basalt interface is variable and uncertain between wells.

To better define the basalt interface in the 200 Area, several seismic reflection and refraction surveys were acquired. These included a historical data set collected in 1979 and 1980 with a total of 80 km of reflection profiles, 15 km of which were recently digitized and reprocessed, 24 km of reflection profiles collected in 2008, and an additional 23 km in 2009 and 2011, all with the primary objective of mapping the basalt interface (Williams et al., 2012a). In addition, vertical seismic velocity profiles were conducted in 29 wells to aid in seismic reflection processing (Peterson, 2010). A comprehensive analysis of these data in addition to interface control points from well data and well-based geologic cross sections was conducted in 2012 to map the basalt interface beneath the 200 Area (Williams et al., 2012a–d). The combined results of that effort over the northern portion of the 200 Area as is shown in Figure 5, including locations of seismic lines and geologic cross sections used in the analysis.

The basalt interface topography shown in Figure 5 provides Hanford Site management and remediation operators with critical information concerning preferred flow pathways for contaminated water to the Columbia River. For example, a large depression in the subsurface basalt boundary is observed north of the 200 Area between Gable Mountain and Gable Butte, bounded to the northwest and southeast by topographic highs that rise above the water table and serve as horizontal flow boundaries (green hachure fill). This depression is known

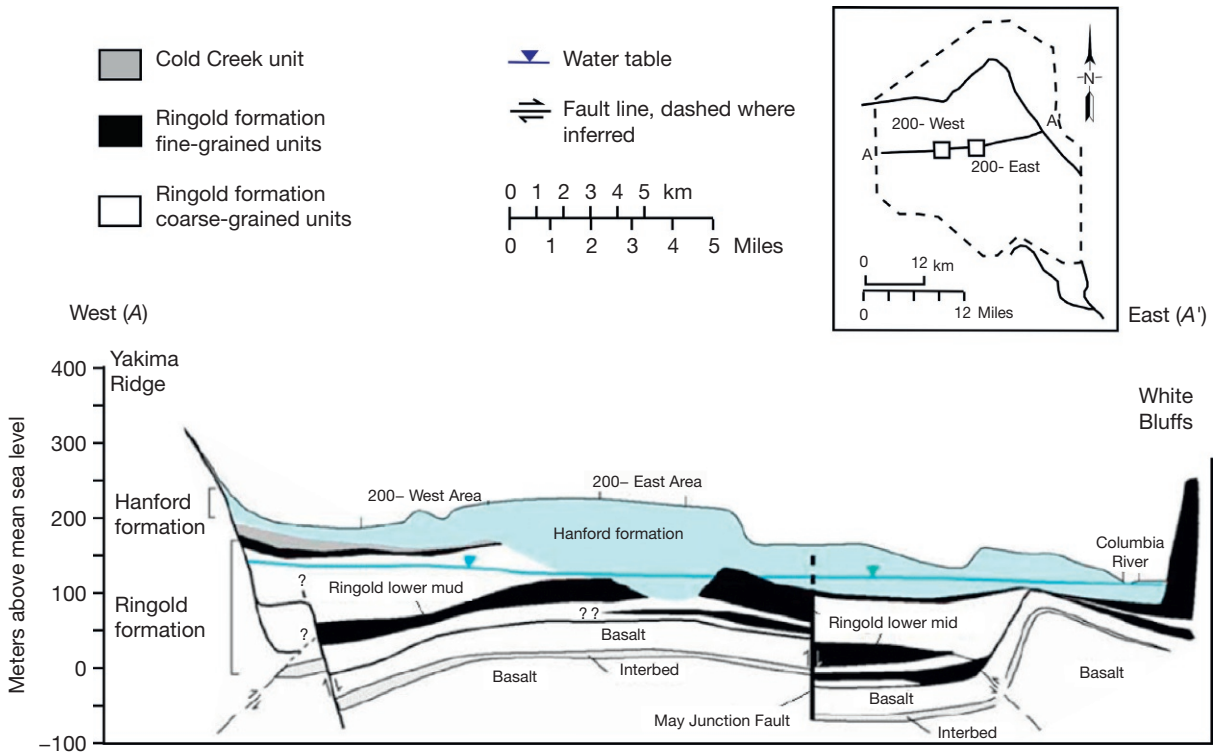


Figure 4 Stratigraphic cross section showing the unconfined aquifer beneath the Hanford 200 Area (DOE/RL, 2012). The primary stratigraphic units for the aquifer are the Hanford formation, consisting of high-permeability sands and gravels, bounded below by low-permeability basalt flows and/or Ringold formation muds. Basalt topography has a strong influence on groundwater flow patterns.

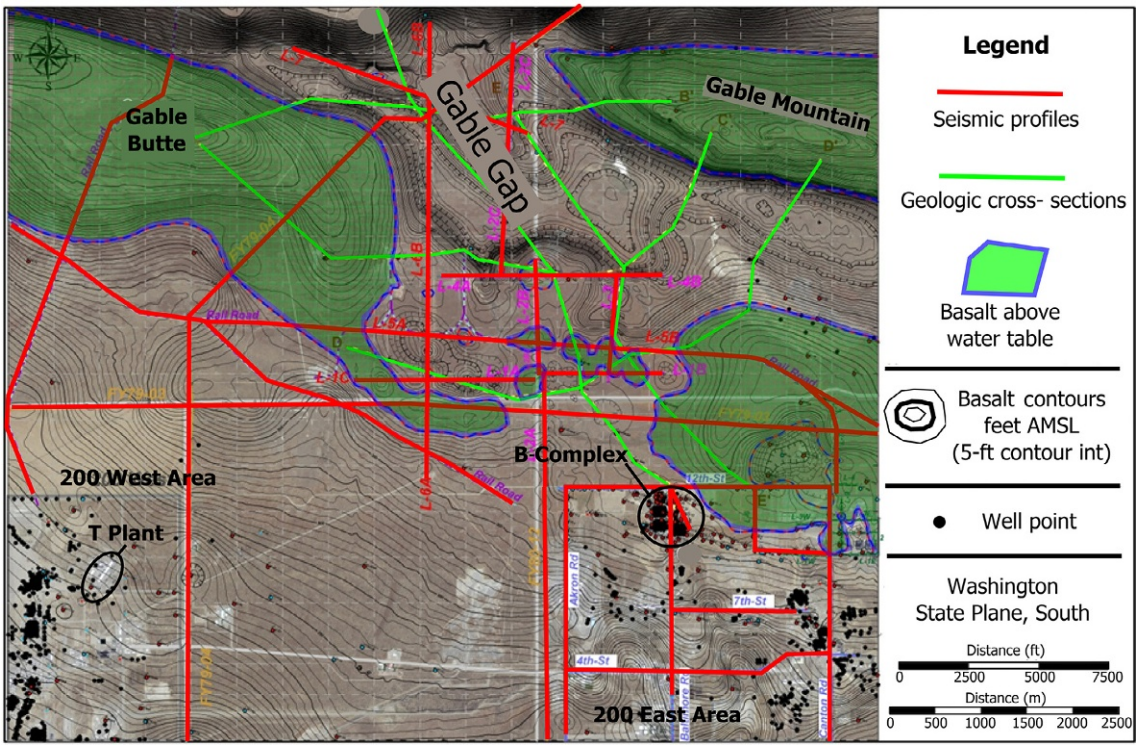


Figure 5 Subsurface basalt interface topography at the Hanford 200 Area inferred from combined seismic reflection, vertical seismic profile, geologic cross section, and well control data (modified from Williams BA, Thompson MD, and Miller SF (2012) Integrated *Surface Geophysical Investigation Results at Liquid Effluent Retention Facility, 200 East Area, Hanford, Washington*. Richland, WA: CH2MHill Plateau Remediation Company). The region between Gable Butte and Gable Mountain (the Gable Gap) does not contain a basalt flow barrier for groundwater flow (shown by green shading) and provides a possible flow path for contaminants originating in the 200 to flow northward toward the Columbia River.

as the Gable Gap and is thought to provide a primary pathway for groundwater originating in the 200 East Area (lower-right corner of Figure 5) to flow north westward through the horizontal flow boundaries associated with Gable Mountain and Gable Butte and onward to the Columbia River. Overall, the reflection data revealed the complex topography of the basalt interface and significantly improved understanding of groundwater flow beneath the 200 Area, enabling operators to more effectively manage current and emerging groundwater contaminant plumes.

11.17.6 Groundwater/River Water Interaction Beneath Hanford 300 Area Infiltration Ponds

The 300 Area, located on the southernmost boundary of the Hanford Site (Figure 1), was the center of Hanford research operations and housed uranium fuel rod fabrication facilities. Liquid waste streams from research and production operations were discharged into infiltration ponds and trenches located on the eastern margin of the 300 Area, adjacent to the Columbia River. The southern ponds and trenches are shown in Figure 6. Between 1943 and the mid-1980s, copper, aluminum, nitrate, and between 33 and 59 metric tons of uranium were discharged into the infiltration ponds (WHC, 1993), creating a groundwater uranium plume that is currently the primary contaminant of concern at the 300 Area.

A 1993 modeling study predicted that groundwater uranium (U) concentrations would decrease to the 30 $\mu g\,l^{-1}$ drinking water standard in 3–10 years (US EPA, 1996; WHC, 1993). A review in 2006, however, showed that, despite source control measures, uranium concentrations remained mostly unchanged (Hartman et al., 2006). Subsequent studies commissioned by the USDOE determined that uranium hot spots within the groundwater vary with seasonal fluctuations in Columbia River stage levels (Williams et al., 2008). During high river stage, groundwater rises into the smear zone. The smear zone is defined herein as the vertical interval of the subsurface that is periodically occupied by the water table. In sediments where high uranium levels exist, uranium desorbs from sediments and flushes into the water phase (Peterson et al., 2008). The highest concentrations of desorbed uranium occur during high stage in locations where water rises into the smear zone beneath former process ponds and trenches where residual uranium exists. Residual sorbed uranium concentrations tend to increase with increasing elevation in the smear zone, with the highest water table events solubilizing more uranium (Williams et al., 2008). In addition to smear zone flushing, Columbia River water intrudes into the high-permeability sand and gravel aquifer during high stage. Columbia River water chemistry further facilitates uranium desorption and thereby enhances uranium release into the aquifer. As stage decreases, groundwater flow again moves toward the river, carrying with it the liberated uranium.

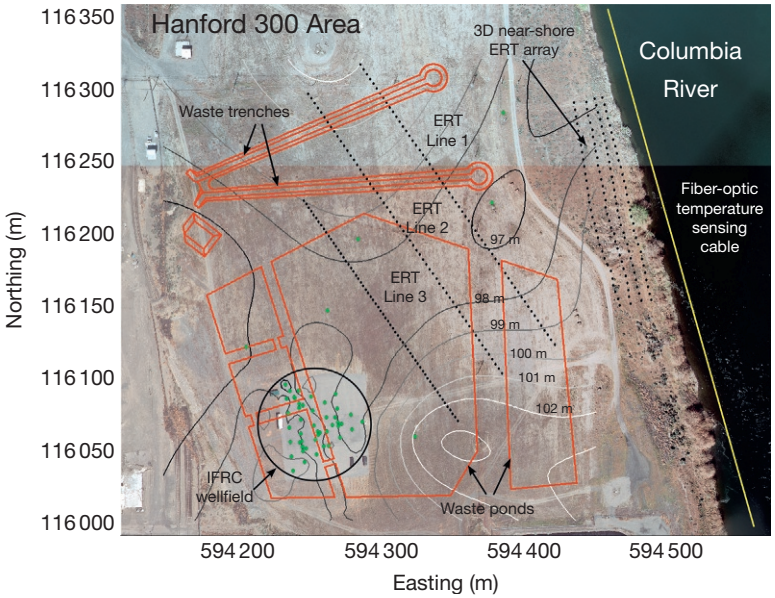


Figure 6 Plan view of Hanford 300 Area showing former waste infiltration ponds and trenches (red), including locations of ERT and temperature arrays designed to understand groundwater/river water interaction within the aquifer, which plays a governing role in uranium desorption and transport to the river.

Given the intimate coupling between groundwater/river water interaction and uranium release, extensive geophysical studies have been undertaken to understand the nature of river water intrusion into the 300 Area. Slater et al. (2010) used waterborne electrical resistivity and induced polarization surveys to map the thickness of the Hanford unconfined aquifer at the river bed, which is indicative of paleochannel structures that cause preferential groundwater flow (**Figure 7**). Capitalizing on the temperature contrast between groundwater and river water, they also used fiber-optic distributed temperature sensing methods to map groundwater discharge along a 1.5 km stretch of shoreline adjacent to the 300 Area, revealing variable, time-dependent, stage-driven groundwater discharge and river water intrusion along the shoreline.

Building on the work of Slater et al., Johnson et al. (2012) installed a three-dimensional (3-D) array of electrodes near the 300 Area shoreline to image the stage-driven inland river water intrusion and retreat from the 300 Area aquifer using 4-D electrical resistivity tomography (see **Chapter 11.08**; **Figure 6**). The objective was to capitalize on the fluid specific conductance contrast between groundwater ($\sim$0.04 S m^{-1}) and river water ($\sim$0.02 S m^{-1}), and the corresponding decrease in bulk conductivity within the aquifer as river water replaces groundwater, to image changes in bulk conductivity associated with river water intrusion and retreat. They collected and inverted 273 surveys over 75 days. Methods for distilling the results using time-series and time-frequency analysis on the river stage and bulk conductivity time series of the elements comprising the inversion mesh were also developed to aid in the understanding of the system. Their analyses revealed spatial and temporal dependence of river water intrusion and retreat on river stage variations as controlled by upstream dam operations. **Figure 8** shows a correlation analysis between stage and bulk conductivity at several depths and validates the location of a dominant groundwater/river water exchange zone along the shoreline previously identified by Slater et al. (2010).

River water is commonly detected 250 m inland to the IFRC well field (**Figure 6**) during spring runoff peak flow. In order to develop and demonstrate the capability to image river water intrusion far inland of the shoreline, three 2-D electrical resistivity tomography (ERT) lines were deployed at increasing distances from the shoreline to monitor intrusion during the 2011 peak flow event. The three surface electrode lines shown in **Figure 6** were operated continuously from May to September 2011, producing three to four time-lapse data sets per day per line and approximately 1200 data sets in total (Wallin et al., 2013).

Inverse analysis of the data revealed the high sensitivity of ERT data to the sharp conductivity contrast at the time-varying water table. The complicating effects of the moving water table boundary, combined with the inaccurate model constraints provided by smoothness constrained inversion across the sharp water table boundary, rendered time-lapse ERT imaging using standard regularization approaches ineffective (see **Chapter 11.08**). Wallin et al. (2013) demonstrated that this problem could be rectified by removing the regularization constraints between neighboring elements that traversed the water table boundary, thereby enabling the inversion to place a sharp conductivity contrast at the water table boundary without penalty. The resulting algorithm provided a time-varying model regularization to accommodate moving the water table boundary, the location of which was known from wellbore pressure measurements.

Example time-lapse inversion results at the rising, peak, and falling limbs of the 2011 Columbia River hydrograph are shown in **Figure 9**, which compares results with and without the water table boundary correction. The results are shown in terms of the change in conductivity from low-stage baseline conditions in early May 2011. River water intrusion is shown

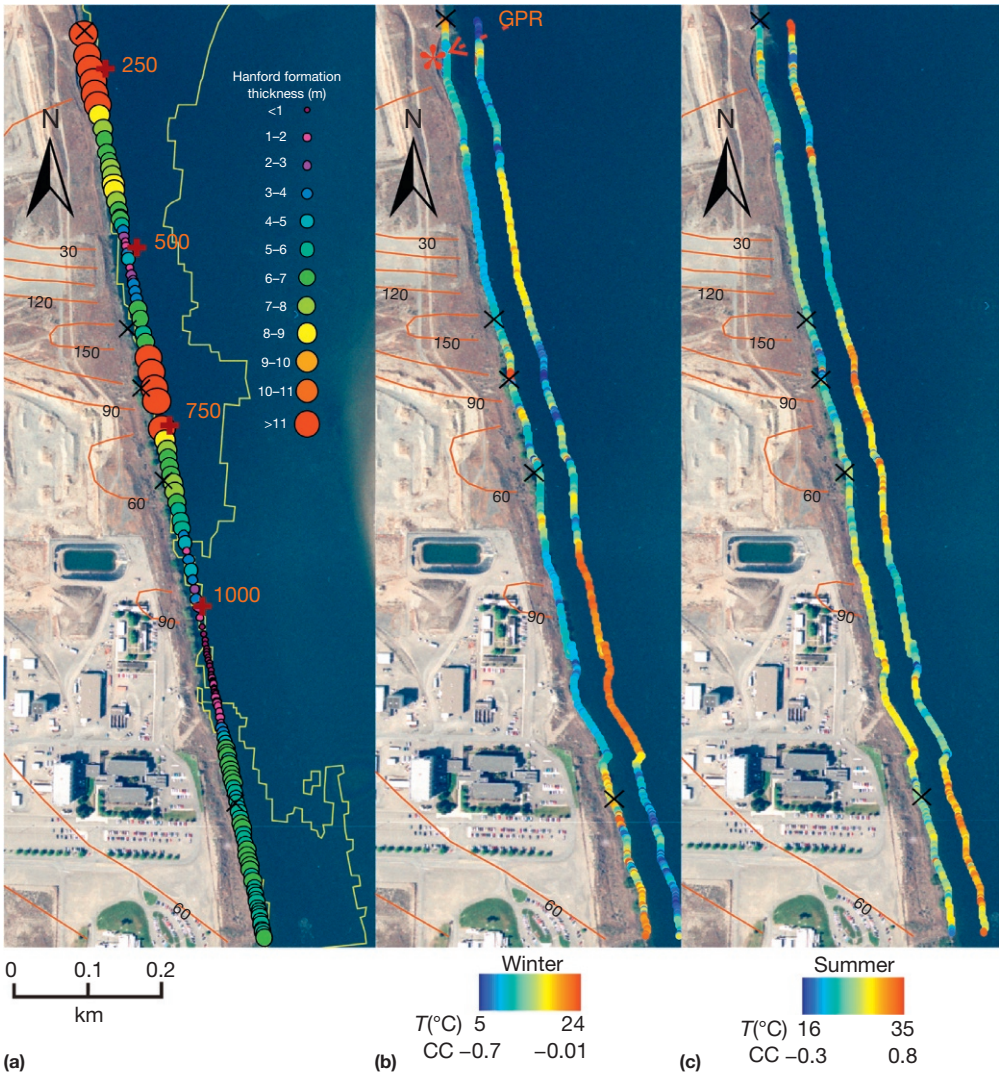


Figure 7 (a) Unconfined aquifer thickness estimated from continuous waterborne electrical imaging measurements. Black crosses show the locations of known springs (Williams et al., 2008). (b) Uranium contours ($\mu g\ l^{-1}$) (Williams et al., 2008) and temperature (T) distribution over 1.5 km of fiber-optic distributed temperature sensing (FO-DTS) for measurements at low stage on 31 March 2009; correlation coefficient between temperature and river stage on 31 March 2009 is also plotted (20 m offset for clarity; T is on the left, CC on the right); asterisk shows the area where a structure interpreted as paleochannel was observed with a GPR survey. (c) T distribution over 1.5 km of FO-DTS for measurements at low stage on 2 August 2009; correlation coefficient between temperature and river stage on 2 August 2009 is also plotted (20 m offset for clarity; T is on the left, CC on the right). Reproduced from Slater LD, Ntarlagiannis D, Day-Lewis FD, et al. (2010) Use of electrical imaging and distributed temperature sensing methods to characterize surface water-groundwater exchange regulating uranium transport at the Hanford 300 Area, Washington. *Water Resources Research* 46.

as a negative change in conductivity (cooler colors). The high conductivity difference band in Figure 9(b) is the variable saturation region where sediments that were unsaturated during low stage have become saturated (and thus more conductive) during high stage. The top of this band denotes the location of the water table at each time step. The vadose zone also shows an increase in conductivity over time, likely caused by warming during the summer months, particularly at the surface. The time-lapse images that account for the water table (Figure 9(b)) display a marked improvement of the standard time-lapse inversion (Figure 9(a)), as validated by wellbore measurements (Wallin et al., 2013).

Building upon the near-shore 3-D and inland 2-D river intrusion monitoring efforts, a 3-D time-lapse monitoring array was installed to monitor river water intrusion during the 2013 spring runoff event (Johnson et al., 2013). An oblique view of the computational mesh showing the locations of electrodes, infiltration galleries, surface features, and the Columbia River channel is shown in Figure 10(a). Mesh surface topography was determined from light detection and ranging data collected at the site. The color scale represents the bulk conductivity at baseline, which was collected 1 April 2013. Figure 10(b) shows four cross sections of the baseline image, which resolves several important features including zones of elevated conductivity associated with vadose zone contamination beneath infiltration trenches; the location of the contact between the Hanford formation, which makes up the primary aquifer; and the Ringold formation, which largely acts as the lower aquifer boundary. The

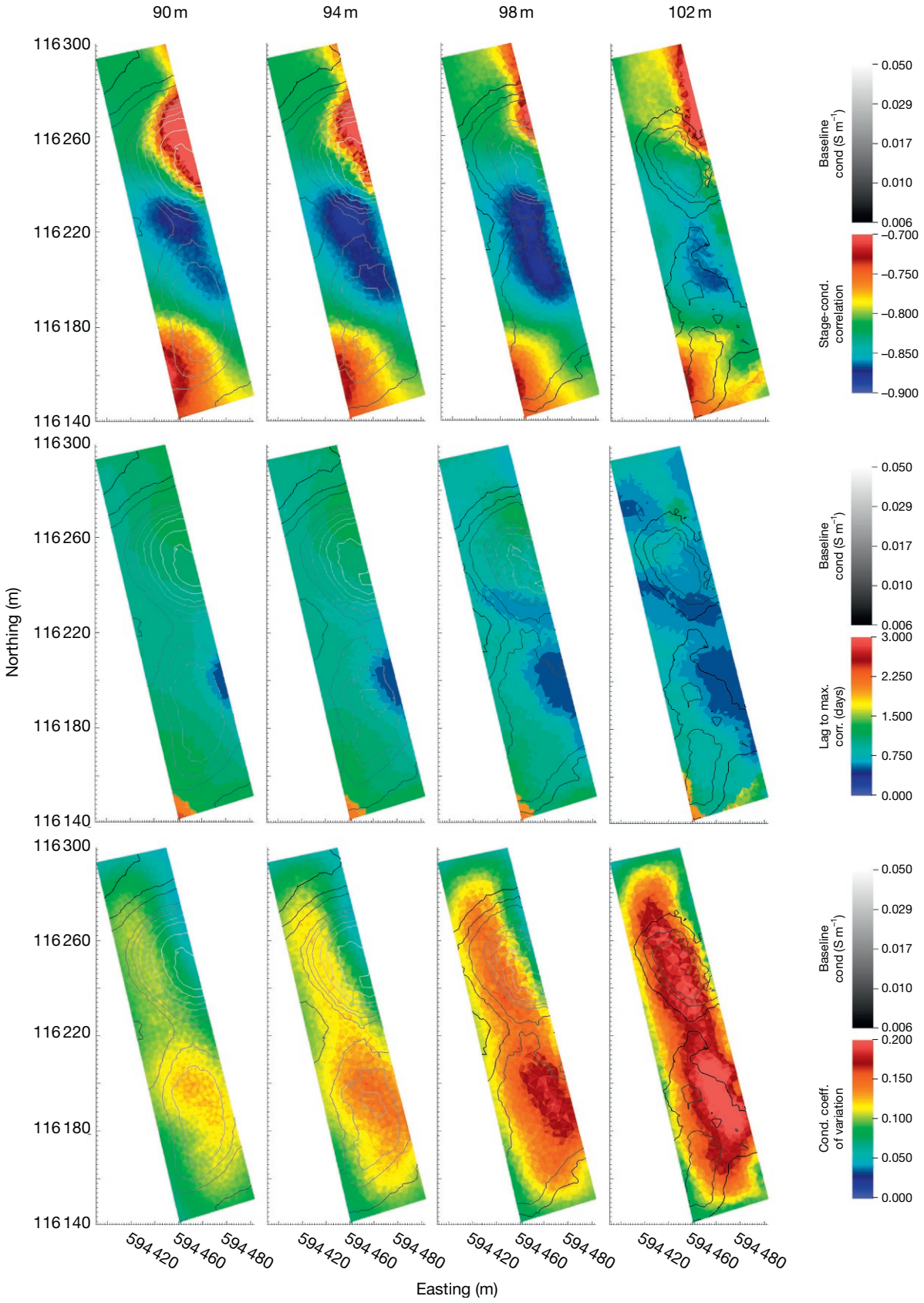


Figure 8 Correlation analysis between bulk conductivity and river stage at four depths beneath the near-shore 3-D ERT array shown in **Figure 6**. Contour lines in each row show the bulk conductivity low near the center of the array indicative of an incised paleochannel. (Top row) Maximum correlation between stage and bulk conductivity time series. As stage rises, river water intrudes, and bulk conductivity decreases, causing the large negative correlation between stage and bulk conductivity in the most active zone (blue color) near the center of the array. (Middle row) Lag time to maximum correlation indicating the time lag between a rise in river stage and a corresponding decrease in bulk conductivity. Shorter lag times are indicative of higher permeability. (Bottom row) Coefficient of variation in the bulk conductivity time series, indicating the relative variation in bulk conductivity with time. Larger variability near the surface suggests more groundwater/river water exchange occurs near the water table than deeper in the formation. Reproduced from Johnson TC, Slater LD, Ntarlagiannis D, Day-Lewis FD, and Elwaseif M (2012) Monitoring groundwater-surface water interaction using time-series and time-frequency analysis of transient three-dimensional electrical resistivity changes. *Water Resources Research* 48.

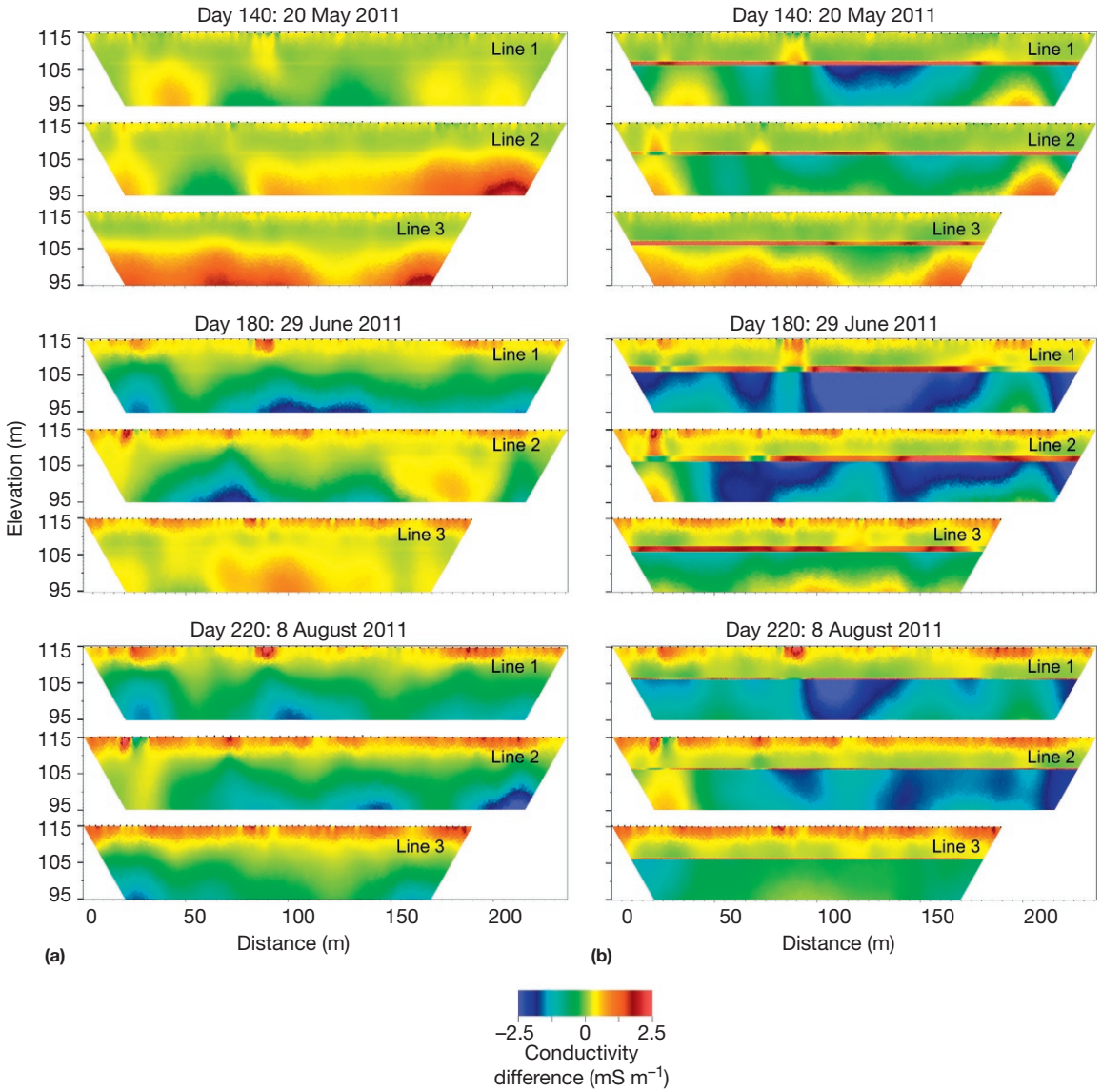


Figure 9 Comparison of time-lapse ERT inversions to image inland river water on lines 1, 2, and 3 (**Figure 6**) using (a) standard smoothness constrained inversion and (b) modified constraints that accommodate the time-varying water table (reproduced from Wallin EL, Johnson TC, Greenwood WJ, and Zachara JM (2013) Imaging high stage river-water intrusion into a contaminated aquifer along a major river corridor using 2-D time-lapse surface electrical resistivity tomography. *Water Resources Research* 49(3): 1693–1708). The full time-lapse sequence (approximately four images per day for 125 days) can be downloaded at http://onlinelibrary.wiley.com/doi/10.1002/wrcr.20119/suppinfo.

shape of the Hanford–Ringold contact is known to exhibit significant control on groundwater movement and uranium transport (Hammond and Lichtner, 2010).

The water table boundary modeling approach applied by Wallin et al. (2013) required a highly refined computational mesh within the smear zone. In terms of the number of mesh elements, that approach required an excessively large mesh when extended to 3-D. Instead, Johnson et al. (2013) (Johnson et al., 2014) devised a warping mesh inversion, whereby the computational mesh warps to the known water table boundary without changing mesh topology, thereby facilitating consistent regularization constraints in the time dimension. For the 3-D application shown here, the warping mesh algorithm reduced the required mesh size from approximately 6.0 million elements to approximately 0.9 million elements, provided a significant increase in computational efficiency. Johnson et al. used the warping mesh algorithm to monitor river water intrusion from May to September 2013, collecting two time-lapse data sets per day, with each survey consisting of approximately 140 000 measurements. In this case, the entire imaging process (i.e., from data collection through inversion and archiving) was completed autonomously in near real time using wireless data transfer and high-performance computing resources, with the time from the beginning of a survey until the time inversion results were available being about 10 h. The real-time imaging provided information concerning the onset of river water intrusion to scientists conducting time-sensitive microbial sampling at the site. Example results using the mesh warping algorithm are shown for two

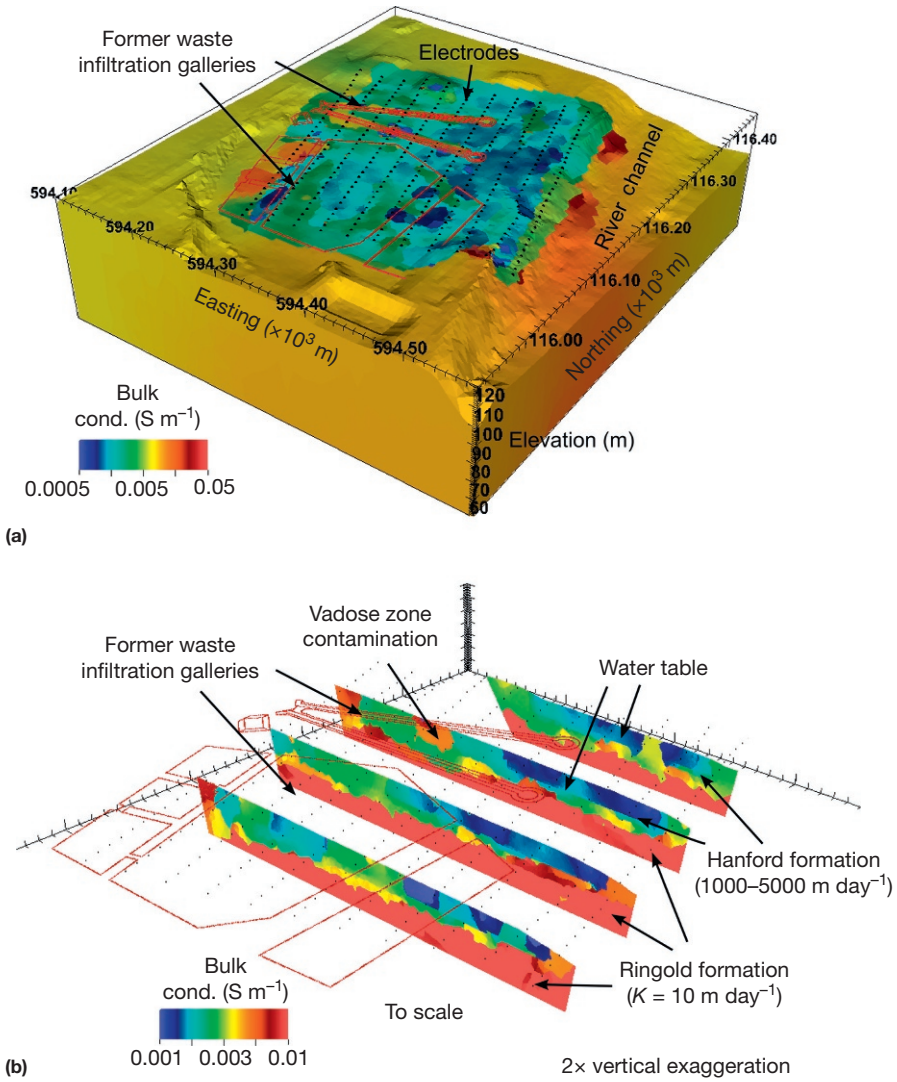


Figure 10 (a) Foreground computational mesh and baseline inversion results for the Hanford 300 Area 3-D time-lapse river intrusion monitoring effort. Note that regions outside of the imaging zone (i.e., beneath the electrodes) are not resolved, but are shown here for topographic relief. (b) Cross sections of baseline inversion results.

different times in Figure 11, one at relatively low stage and one at relatively high stage. Overall, the 3-D ERT monitoring provided a significantly increased understanding of groundwater movement in the 300 Area in comparison to the 2-D monitoring. For example, the shape of the Hanford–Ringold contact provided in the baseline inversion is consistent with the dominant zone of intrusion revealed from the time-lapse inversions. Importantly, while providing additional information, these results are consistent with the previous efforts of Slater et al. (2010), Johnson et al. (2012), and Wallin et al. (2013).

11.17.7 Vadose Zone Contaminant Mapping at the Hanford 200 East Area B-Complex

The B-Complex, located in the Hanford 200 East Area, is a former waste disposal facility consisting of 40 single-shell HLW storage tanks, many of which have leaked or experienced accidental overfill episodes, and several outlying subsurface infiltration galleries. The waste streams introduced into the vadose zone were highly saline and created contaminated zones of elevated bulk electrical conductivity in contrast to the native low conductivity sands and gravels (Lindenmeier et al., 2002). The waste poses a significant risk to groundwater quality, and determining the distribution of vadose zone contamination remains one of the most significant challenges limiting remediation and closure of Hanford Site waste disposal facilities.

Given the electrical conductivity contrast between contaminated and pristine soils at the B-Complex, a large-scale surface resistivity survey was conducted with the intent of using ERT to map the distribution of subsurface contamination resulting from past waste disposal operations (Rucker et al., 2007). Figure 12(a) shows the relative locations of major B-Complex features, including the tank farms, infiltration galleries, wells, and surface electrodes used in the survey. Stainless steel

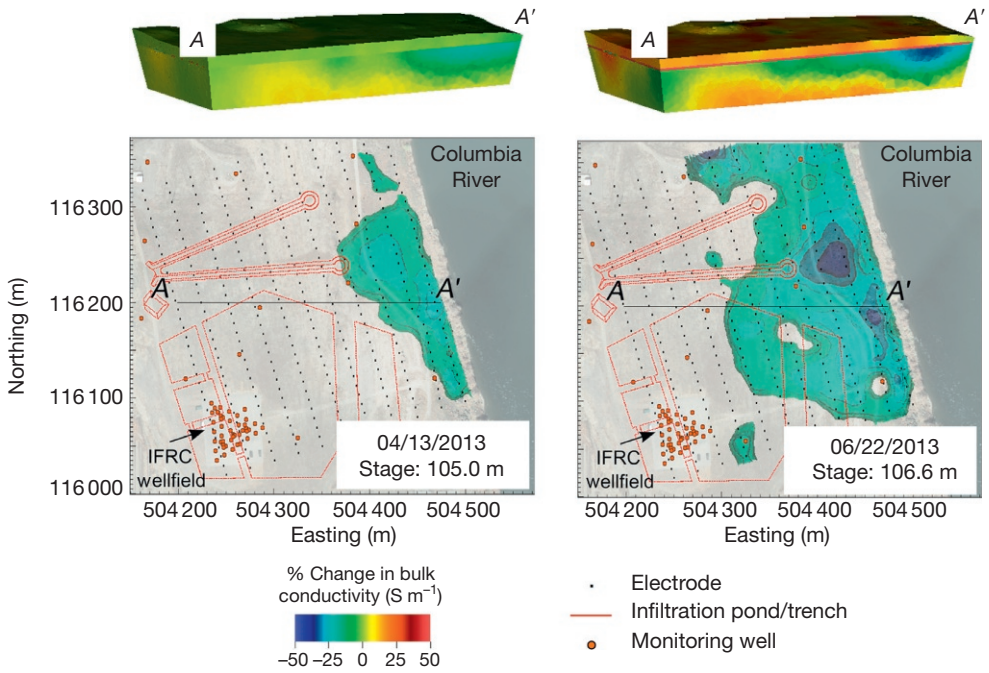


Figure 11 Examples of 3-D river water intrusion images collected at relatively low stage on 13 March 2013 and high stage on 22 June 2013. Upper images show cross sections along $A-A'$, and lower images show several transparent isosurfaces of change in conductivity superimposed on a transparent aerial image of the 300 Area. Negative changes (cooler colors) denote greater river water concentration.

electrodes were laid out in lines parallel and perpendicular to the tanks at a separation of ~30 m with an intraelectrode spacing of 6 and 3 m inside and outside of the tank farm boundaries, respectively. The 36 lines across the site varied in length from 417 to 816 m. Remote electrodes were placed orthogonally to the northeast and northwest ~1900 m from the center of the domain in either direction. The final data set consisted of 208 411 pole–pole measurements (see Chapter 11.08) collected on 4859 electrodes.

Figure 12(b) shows the locations of the primary waste releases into B-Complex infiltration galleries (Williams, 1996). The area of the circles represents the volume of waste released in millions of liters (ML), and the color scale represents the approximate ionic strength of the waste. Regions where more than 18.9 ML were released are indicated in text. Higher disposal volumes were generally accompanied by more dilute waste. The exception occurred in the BY-Cribs Area, which was subjected to high waste volumes with high ionic strengths. Note that Figure 12(b) does not include unplanned releases that occurred within the tank farm boundaries. Unplanned releases are accidental spills, leaks, or tank overflow events that, in comparison to planned releases, are smaller in volume but higher in contaminant concentration.

Figure 13 shows a plan view of the ERT inversion results over the entire B-Complex in terms of transparent bulk conductivity isosurfaces (Johnson and Wellman, 2013; Johnson et al., 2010). These isosurfaces outline the 3-D extent of elevated bulk conductivity regions that are diagnostic of vadose zone pore water contamination beneath the BX-Trenches, BY-Cribs, B7 and B8 Cribs, and the Tile Field shown in Figure 12(a). Elevated conductivities within the tank farm boundaries are primarily dominated by the dense arrays of metallic monitoring wells and tanks. However, the high conductivity regions beneath the waste diversion boxes within the tank farm boundaries are well below the depth of metallic transfer pipes and, given the magnitude of the conductive anomalies, are likely associated with highly contaminated pore water.

According to the waste release inventories shown in Figure 12, the BY-Cribs were subjected to high release volumes at elevated concentrations. The majority of these releases were derived directly from tank wastes and were executed over a 2-year period from 1954 to 1955 in an effort to increase tank storage capacity (DOE/RL, 1999). Consequently, the BY-Cribs are the most contaminated infiltration galleries within the B-Complex. The discharge inventory into the BY-Cribs is consistent with the ERT inversion results, which show bulk conductivities ranging over approximately four orders of magnitude. Figure 13 shows an oblique 3-D view of the ERT inversion results in the BY-Cribs Area. The ERT images suggest that, in addition to vertical migration beneath the cribs, wastes released into the cribs migrated laterally ~100 m in one eastward trending plume and also in a second southeast trending plume. The eastward trending lobe appears to originate from the northeastern-most crib and is oriented at an angle of ~30° from horizontal downward to the vadose zone. In contrast, the southeast trending plume appears to originate at depth from the main plume directly beneath the cribs and is oriented horizontally between 130 and 160 m in elevation. The horizontal orientation of this lobe may be caused by a vertical flow barrier at depth as has been observed to cause perched water in other parts of the B-Complex. The main plume, which is oriented directly beneath the cribs, appears to have migrated vertically, merging at depth from more discrete plumes

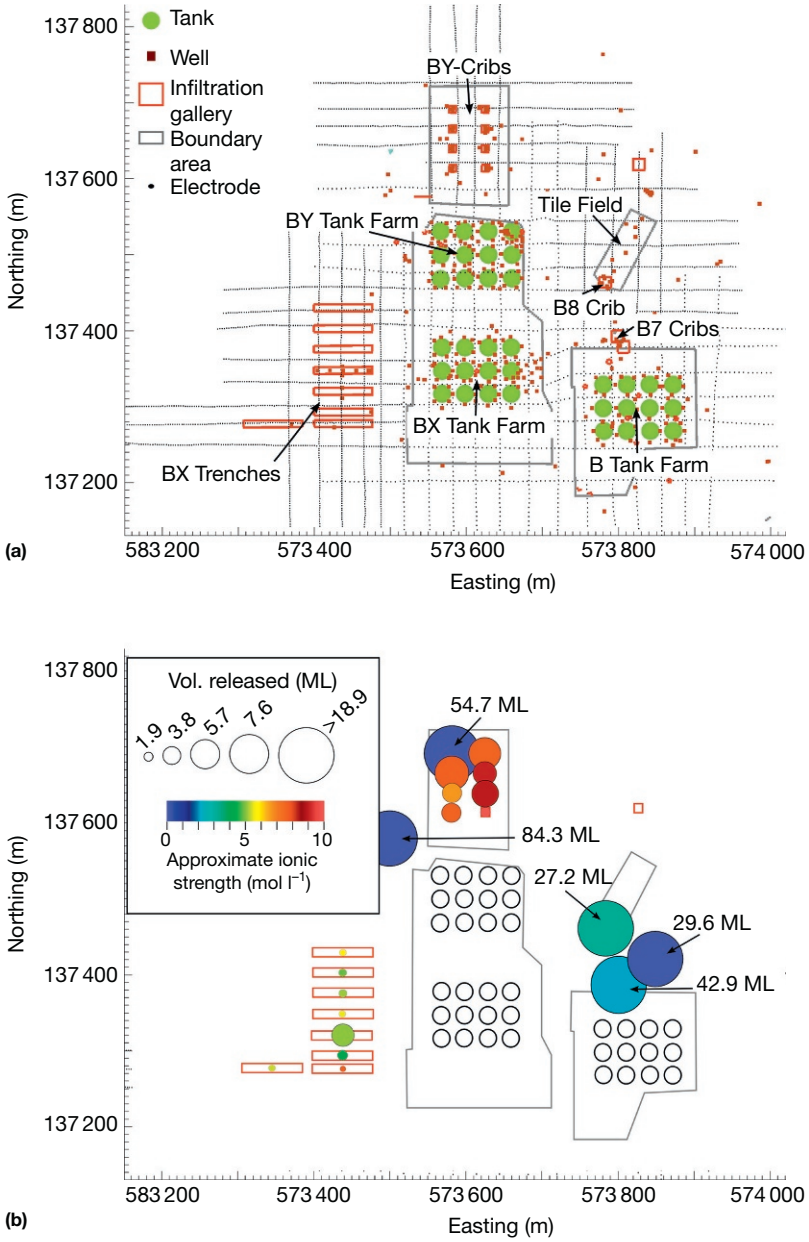


Figure 12 (a) Plan view of Hanford B-Complex showing locations of storage tanks, infiltration galleries, wells, and electrodes. (b) Summary of the major disposal inventories within B-Complex infiltration galleries in terms of volume in millions of liters and approximate ionic strength. Unplanned waste releases (i.e., leaks) within the tank farms are not represented.

occurring shallow in the section. These observations are all consistent with the disposal history shown in Figure 12(b) and have armed site remediation operators with information concerning contaminant distribution and the origin of contaminated groundwater plumes emerging from the B-Complex vadose zone (Figure 14).

11.17.8 Long Electrode ERT Imaging of HLW Inside Tank Farms

Disposal facilities on the periphery of the tank farms are far from the dense network of infrastructure necessary to move and store waste inside the boundaries of tank farms. The electrically conductive anomalies beneath the BX-Trenches, BY-Cribs, and the eastern B7 and B8 Cribs all show finite boundaries that align well with known disposal practices and ancillary hydrogeologic evidence (Figure 13). Within a tank farm, however, there is extensive metallic infrastructure including pipes, tanks, wells, electrical lines, distribution manifolds, and fences. The conductive elements greatly complicate the use of geophysical methods that aim to estimate electrical properties of the subsurface. Specifically, the infrastructure is orders of magnitude more conductive than the waste being imaged and applied currents can be concentrated along these features, reducing sensitivity to surrounding materials. These

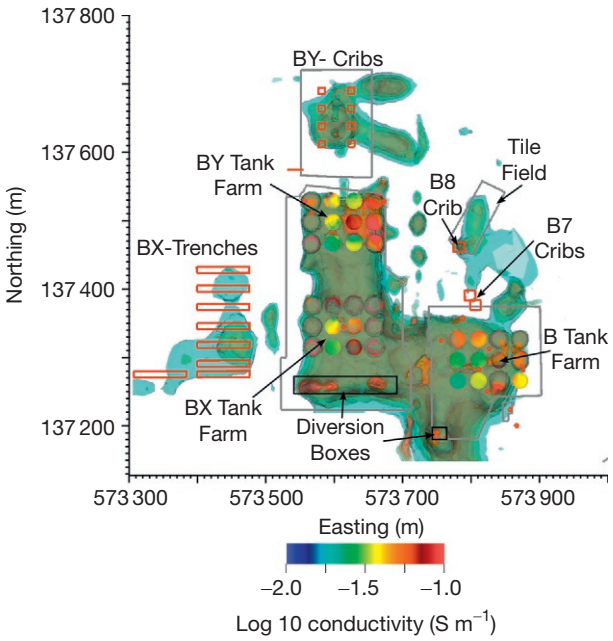


Figure 13 Plan view of 3-D bulk conductivity distribution beneath the B-Complex with elevated regions outlined by transparent isosurfaces. Elevated bulk conductivity regions originating from the infiltration galleries outside of the tank farms are diagnostic of high ionic strength (e.g., contaminated) pore water in the vadose zone. Bulk conductivity within the tank farm boundaries is largely dominated by the metallic tanks and dry wells and is therefore inconclusive concerning contaminant distribution. The exception to this includes the regions beneath the diversion boxes, where deep metallic infrastructure is minimal and elevated conductivities are indicative of vadose zone contamination.

complications are especially difficult for methods that rely on inversion because most algorithms do not account for current focusing, scattering, and secondary voltages. Referring again to Figure 13, the conductive anomaly imaged using surface electrodes within the B-Complex tank farm appears to mimic the tank farm footprint with little differentiation inside the boundary, limiting the ability to delineate or quantify past releases from specific tanks.

Given the strong influence of metallic infrastructure on electrical resistivity response, it may be advantageous to incorporate the infrastructure directly into the acquisition and modeling approaches. The steel-cased wells originally used for drywell geophysical logging, for example, can be used as long electrodes in an electrical resistivity survey. It has been shown that resistivity models developed from long electrodes will have lower resolution than that provided by vertical arrays of discrete electrodes in the same location because the voltage potential is nearly constant along the length of the long electrode. The spatial distribution of electrical resistivity, as calculated by inversion models, is more highly resolved in areas of high voltage gradients. However, in many cases, the metallic infrastructure provides the only option for noninvasive plume mapping within a tank farm. Using long electrodes for resistivity imaging is also an attractive option for imaging due to their large numbers at Hanford and being uniformly distributed within many of the tank farms. In addition, due to the installation methods and the long time that the casings have been in the ground, long electrodes are well grounded and electrically isolated from other metallic infrastructure. Finally, the wells are easily accessible with known location and dimensions. Horizontal drywells and piping near tanks also exist in sufficient number to be attractive for use. However, there has not been an opportunity to implement this additional infrastructure yet, due to limited access and lack of isolation.

Tank farms can have 100 or more steel-cased wells, completed in the vadose zone either as drywells for logging (from 10 to 45 m below ground surface) or as groundwater sampling wells (over 100 m deep). In all cases, the wells are modeled as a series of extremely electrically conductive cells. In addition, the forward model and the inverse model grids are decoupled so that the conductive long electrodes are not part of the inverse model solution. The decoupling ensures that the effects of the conductive well do not interfere with the algorithm's ability to resolve lower conductivity contaminant targets due to regularization. The algorithms have been benchmarked against analytical solutions using a mixture of point and long electrodes with minimal deviations. The methodology has also gone through some validation using a pilot-scale field experiment (e.g., Rucker et al., 2012).

The use of long electrodes for understanding waste distribution beneath tanks has a rich history at Hanford. Key (1977) tested wells by measuring the change in the potential field generated from a cathodic protection system that inhibited corrosion. The tests were meant to evaluate external tank and pipeline leak detection systems and it was found that the wells were sufficiently sensitive to measure small changes in conductivity from a simulated leak. Ramirez et al. (1996) and Daily et al. (2004) discussed using both wells and tanks as electrodes for imaging conductive anomalies beneath a ¾-scale mock tank outside a tank farm. For these investigations, long electrode analogues were fashioned by shorting a series of point electrodes within a borehole and showed promising results with respect to detecting the emergence of an electrically conductive anomaly. More recently, Rucker et al. (2010, 2013) had applied the long electrode ERT (LEERT) methodology within an actual tank farm to image suspected releases from tanks.

As a demonstration, the LEERT method is applied within the SX Tank Farm, located in the 200 West Area of Hanford, to understand the contaminant distribution from historical leaks. The SX Tank Farm consists of 15 single-shell tanks with a capacity of 3.8 ML per tank for a total of 56.8 ML (Knepp, 2002). The tanks received waste generated from the reduction–oxidation process used to extract plutonium from fuel rods in the 1950s and 1960s. Seven tanks in the SX Tank Farm are known or assumed to have leaked (Johnson and Field, 2010). To reduce the risk of future leaks, all drainable liquid in each of the single-shell tanks was removed and transferred to double-shell tanks (DOE/ORP, 2005). Some tank farms presently have interim surface barriers (ISB) designed to reduce the amount of surface recharge into the vadose zone, thereby slowing the migration of subsurface contaminants to groundwater. The SX farm is one of the next candidates to receive an ISB and LEERT is being used to assess the minimum size of the barrier by evaluating the lateral extent of the plumes.

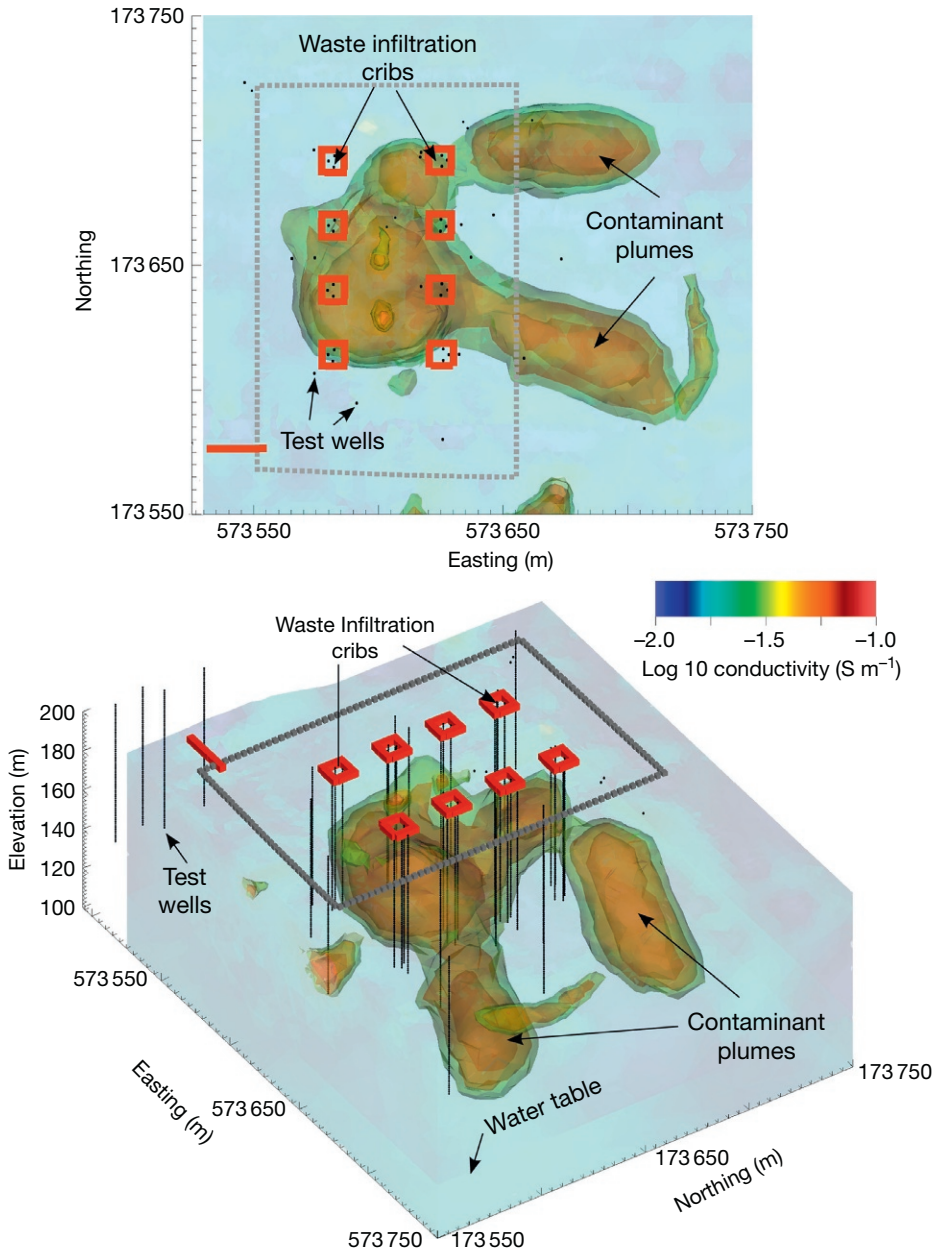


Figure 14 Oblique view of BY-Cribs Area surface ERT inversion results. Color map is the same as that provided in **Figure 13**.

Figure 15 shows the suspected or known historically leaking tanks with the associated volume loss values based on recently revised estimates (Johnson and Field, 2010; Rogers, 2013). To aid in the understanding of consequences related to the leaks, soil samples for geochemical analyses were collected inside and outside the tank farm during a direct push drilling campaign. Results of the geochemical analyses are also displayed in Figure 15. Soil nitrate values are displayed as disks of varying sizes representing specific ranges of logarithmically transformed concentrations. The geochemical data show that within the tank farm, the highest nitrate concentrations are nearest tank SX-108, which had the largest suspected leak. The values near the tank were upward of 45 000 $\mu g\ g^{-1}$ of soil. On the outskirts of the tank farm and outside the LEERT modeling domain, a secondary nitrate plume exists. This was likely caused by nearby disposal cribs or other leaky infrastructure.

The 3D LEERT survey of SX Tank Farm used 96 wells, with well spacing varying between 5 and 168 m. Data acquisition included 8230 pole–pole measurements with reciprocals included. Reciprocal measurements are conducted by switching the current transmission and voltage potential dipoles to obtain complementary pairs of voltage data and can be useful in determining the quality of a resistivity data set. In theory, the reciprocal measurements should be nearly identical to the

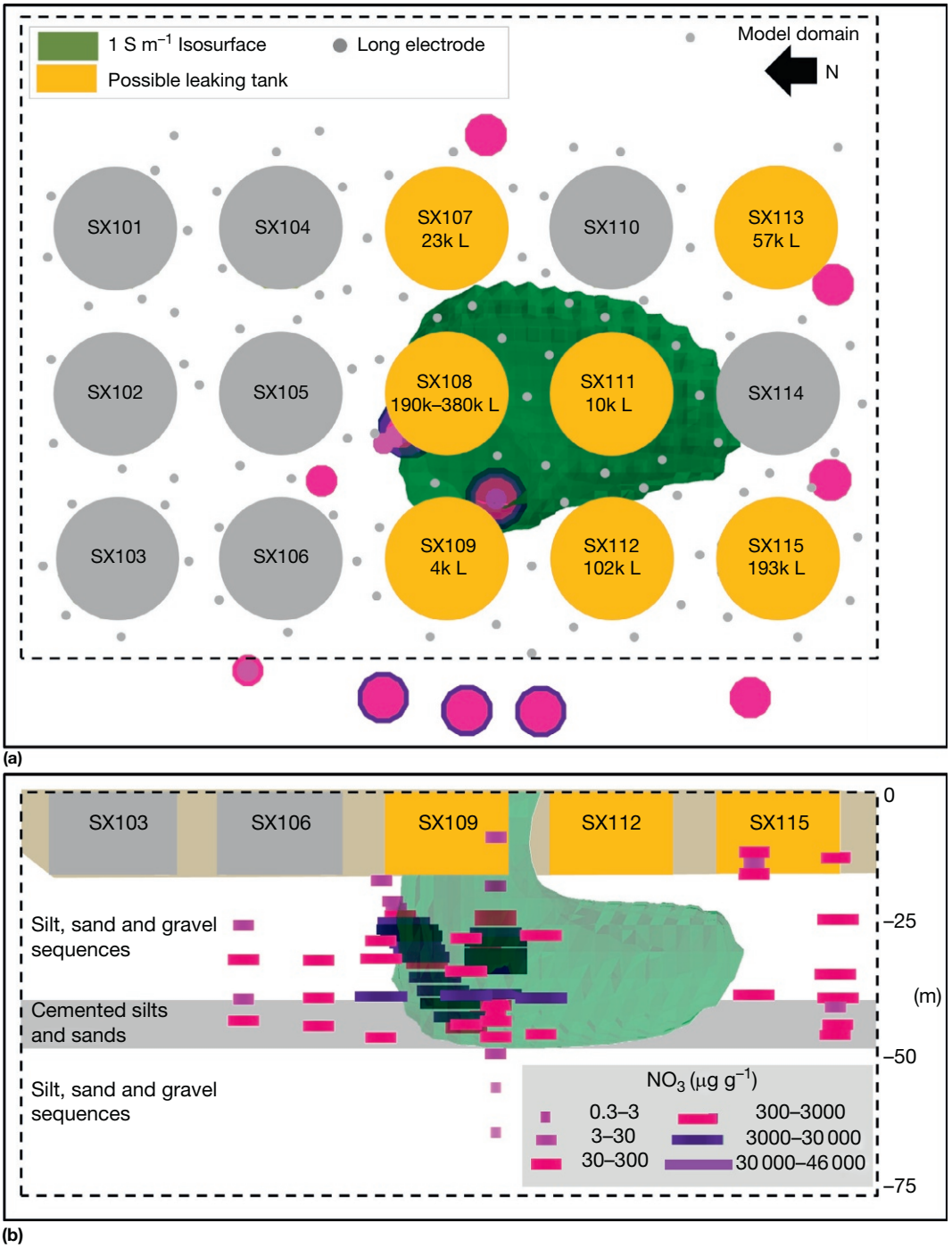


Figure 15 Long electrode ERT results at 1 Ωm with nitrate (NO3) in soil analytical results and known geologic stratigraphy. Nitrate results are displayed in classed posts with relative size indicative of an order of magnitude of nitrate concentration in μg g^{-1}. Note the effect of the cemented layer on vertical plume migration. Part (a) is plan view and part (b) is cross section (Johnson and Field, 2010; Rogers, 2013).

forward measurements and those data that differ significantly provide a justification for their removal. As such, a 5% difference between reciprocal apparent resistivity values was used as a means to remove pairs of bad data. The steel wells were modeled with resistivities of 0.006 Ωm (or conductivities of 167 S m^{-1}), with a diameter of 0.1 m, and were assumed to be vertically oriented. The transition of the high conductivity well to the low conductivity soil is buffered by a few layers of intermediate conductivity values in the forward model to accommodate problems typical of numerical models. An abrupt transition in resistivity values over a short interval will cause numerical instabilities and a misrepresentation of

current flow across the model cells. Additionally, given the short length of the well, the unconsolidated nature of the Hanford formation, and the method of drilling (cable tool), the vertical well orientation is likely a safe assumption. Based on limited analyses, a typical deviation of 10° for wells less than 70 m causes little trouble for the method.

Results of the LEERT inversion model are shown in Figure 15 as a green solid-rendered isosurface that encloses areas with electrical resistivities equal to or less than 1 Ω m (1 S m^{-1}) to focus specifically on the high conductivity target resulting from the spill. The locations of the highest concentrations of nitrate align well with the electrically conductive target inside the tank farm. Also of interest is the shape of the conductive plume. In particular, the plume is elongated in a direction that is consistent with the Cold Creek stratigraphic unit, which dips slightly to the south at a depth of ~38 m. The Cold Creek is a Pliocene-age calcified paleosol layer with cemented sands and silts that extends across most of the 200 West Area beneath the Hanford tank farms. The LEERT results suggest that this stratigraphic unit may be restricting the vertical infiltration of the nitrate plume to groundwater.

Comparing the colocated LEERT inversion results to pore water nitrate from the direct push samples (Figure 16(a)) reveals a reasonable correlation, especially considering the different scales of measurement and the low vertical resolution of the LEERT method. Figure 16(a) shows a positive linear relationship between bulk LEERT conductivity and log pore water nitrate concentration. Rucker et al. (2012) also presented scatter plots for other tank farms where the fit is much worse than shown in Figure 16(a).

Lastly, Figure 16(b) demonstrates a strong power relationship with a correlation coefficient of 0.82 between technetium-99 (Tc-99) and nitrate concentrations in the soil analytical results acquired throughout the SX Tank Farm (Glaser, 2013). Tc-99 is highly mobile, and like nitrate, it moves through Hanford soils relatively unretarded. Tc-99 is also a long-lived beta emitting radioactive isotope and, as such, is one of the contaminants of greatest risk to human health at the site. Unfortunately, there are presently no tools for remotely locating or mapping subsurface distribution of Tc-99, with the exception of extensive direct soil sampling. However, by effectively acquiring electrical measurements of the subsurface and noting the relationships observed in Figure 16(b), variations in Tc-99 can be estimated. Given that both technetium and bulk conductivity are linearly correlated with nitrate, bulk conductivity can potentially be used to infer variations in technetium concentration.

11.17.9 HLW Tank Leak Detection Using Electrical Methods

Efforts are under way to retrieve HLW from compromised single-shell tanks and transfer it to double-shell tanks for interim storage. Ultimately, all HLW will be placed into a stable waste form and disposed in long-term storage facilities (DOE/RL, 2013). Monitoring is conducted during retrieval to detect possible leaks caused by the retrieval activities. Due to limited access to the subsurface in the tank farm, the large tank dimensions, and the dangers posed by drilling near tanks, direct detection of leaks outside of the tanks by sensor emplacement was determined to be infeasible.

The tank waste is highly saline and electrically conductive relative to the native low conductivity sands and gravels of the Hanford formation. Therefore, remote detection using electrical resistivity measurements was developed, tested, and approved as a viable method of detecting tank leaks (Barnett et al., 2002, 2003; Bratton and Boger, 2002; Key, 1977; Reich, 2004; Schofield, 2006). The approach involves using the

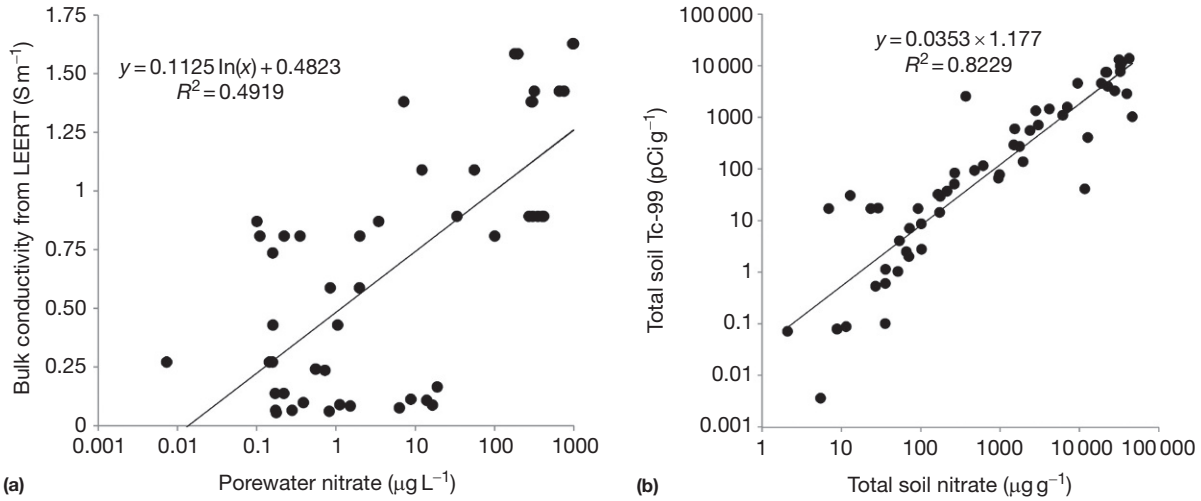


Figure 16 (a) Regression and correlation between bulk conductivity obtained from long electrode ERT and calculated pore water nitrate from total soil nitrate. (b) The relationship between technetium-99 and nitrate soil analytical results at SX farm (Glaser, 2013). Presently, there are no tools that can remotely measure the high-risk technetium distribution in the soil. ERT provides an opportunity to locate high conductivity contaminants such as nitrates and to infer the presence of associated technetium.

drywell casings as current injection electrodes in a pole–pole configuration and using both tanks and drywells as potential measurement electrodes. In some instances, point electrodes are used on the surface or in nested borehole arrays to supplement acquisition in regions that do not have many wells present. The potentials generated using many electrode pairs around a particular tank are monitored over time, and changes in the voltage are evaluated for a possibility of a leak without the need for full-scale tomographic inversion. In this way, the data can be evaluated instantaneously, thus reducing response time for mitigation efforts.

Figure 17 shows an example of the electrical potential field during current transmission on a local long electrode pole near three electrically resistive tanks. The elongated electrode causes the near field to have only a minor vertical gradient. The far field, however, shows more of the hemispherical shape typical of point electrodes. A conductive waste leak from the side or bottom of the tank will change the field and the change will be much more noticeable on those electrodes nearest the leak.

Electrical conductivity depends on many factors (water content, salinity, temperature, and surface charge). Therefore, it is rare that measured electrical conductivity can be used directly to identify a leak. Rather, time series and spatial patterns of electrical conductivity change must be interpreted to identify anomalies that may indicate a leak. To provide rapid leak detection, a time series of measured voltage for subsets of the entire data set is analyzed at each measurement time to look for behaviors that are inconsistent with trends and/or with the behavior of other voltage time series. When an anomalous pattern is identified, this suggests the possibility of a leak, but these data alone do not automatically provide a binary (yes or no) indication of a leak (Calendine et al., 2011; Glaser et al., 2008). To avoid false-positive leak indications, other possible causes are investigated as possible sources of the anomalous response. If no suitable justification can be determined for the observed voltage pattern, then a leak potential assessment process is entered.

Factors that can cause anomalous voltage responses fall into one of four categories: a tank release, environmental changes, anthropogenic perturbations, or system malfunctions. Environmental factors such as diurnal and seasonal temperature changes, local precipitation, and lightning can cause changes in measured voltage. Diurnal temperature effects are usually subtle, with amplitudes that are considerably lower than an actual leak response. Precipitation, on the other hand, can create larger responses, which can be more difficult to differentiate from a leak. Precipitation impacts are most noticeable when using surface electrodes that have high sensitivity near the ground surface. Lightning can cause serious harm to the electrical components of the acquisition system if the strike hits within a short range of the electrode array (including the remote poles). Anthropogenic impacts are caused by changes to onsite electrical equipment used for retrieval (i.e., change in grounding conditions), typically causing voltage shifts or offsets by up to a few millivolts in all data pairs. System malfunction stems from individual electrical components within the data acquisition system failing, again causing a shift in electrical potential data. Leak signatures in the voltage record are more gradual, thus eliminating sudden shifts as a possibility of a leaky tank.

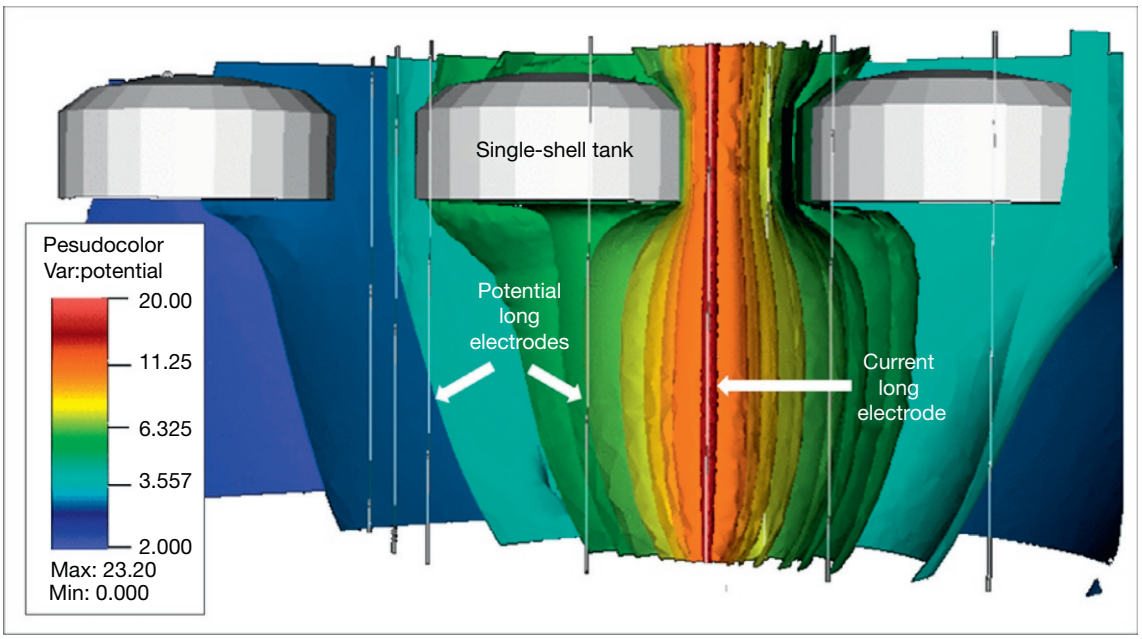


Figure 17 Electrical potential distribution (voltage) generated from current transmission on a long electrode in a tank farm, assuming infinite conductivity and electrically isolated wells and tanks. Reproduced from Johnson TC, Wellman DM, and Triplett M (2014) Modeling of electrical resistivity data in the presence of electrically conductive well casings and waste storage tanks. In *Waste Management Symposia*. Phoenix, AZ: Waste Management.

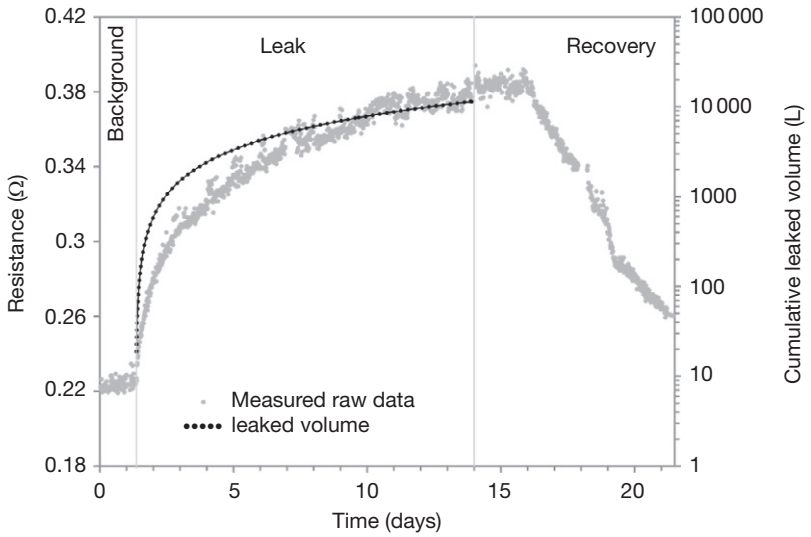


Figure 18 Example data set obtained during a leak injection test on a well near tank S-102 conducted in January 2006. The test was performed to determine the sensitivity of an electrical monitoring system in a fully operating tank farm.

The resistivity-based monitoring approach was first applied commercially in the S Tank Farm, around tank S-102. The system operated for approximately 3 years from 2004 to 2007, during which time a leak injection test was performed to ensure operability and sufficient sensitivity to provide leak detection in a complex setting (Rucker et al., 2007). The leak injection test included ten small leak events over a 4-month period, whereby fluids with conductivity similar to tank waste were injected into a vadose zone well to simulate a tank leak. Resistance data were collected in near real time and analyzed on a once-daily basis to identify data trends indicative of a leak. An assessment was completed at the end of each leak event to evaluate the performance of the technology. As an example, Figure 18 shows the measured resistance relative to cumulative leaked volume for the first leak. Prior to the leak, a static baseline resistance was recorded for approximately 2 years. The figure shows that the onset of the leak, which lasted nearly 13 days at a rate of about 0.67 l min^{-1}, caused a large change in resistance. The polarity of the change, whether voltage increased or decreased, depended on the electrode geometry and its proximity to the leak. Statistical analysis of final test results with all ten leaks indicated that for the test geometry and soil conditions present, the system was capable of detecting a leak of ~6000 l with 95% reliability. It is believed that the system threshold is significantly below this statistical value based on individual leak performances. Figure 18, for example, shows a significant change in voltage within the first few days to a leak of less than 500 l. It should be noted that the leak injection test was performed in parallel with neutron moisture logging of the drywells in the vicinity of the injection well. Neutron logging and gamma logging were the primary ex-tank means of leak detection during waste retrieval operations prior to the implementation of the resistivity-based leak detection system. However, the neutron logs did not indicate a significant variation from background moisture content at any time during the injection of more than 50 000 l (Schofield, 2006).

11.17.10 Engineered Vadose Zone Desiccation Monitoring in the Hanford BC Cribs Area

At the Hanford Site, in situ vadose zone remediation is a potential approach for addressing inorganic and radionuclide contamination located at depths below the limit of direct exposure where remediation is focused on groundwater protection (Dresel et al., 2008). In 2008, the USDOE initiated a treatability test program to evaluate potential deep vadose zone remedies for protection of groundwater (DOE/RL, 2008), focusing on options to mitigate the vertical movement of contaminants toward the water table. To assess efficacy, these remedies require effective methods of performance monitoring in the vadose zone. Engineered soil desiccation was investigated as a potential vadose zone remediation technology, including laboratory studies (Oostrom et al., 2009, 2012a,b; Truex et al., 2011), modeling studies (Truex et al., 2011), and field testing (Truex et al., 2012) conducted as part of treatability test efforts at the Hanford Site.

Desiccating a portion of the vadose zone has the potential to minimize migration of contaminants toward the water table by decreasing the permeability of desiccated sediments. To apply desiccation, a dry gas is injected into the subsurface. The dry gas evaporates water from the porous medium until the gas reaches 100% relative humidity and can no longer evaporate water. Evaporation can remove pore water and result in very low moisture content within the desiccated zone (Oostrom et al., 2009; Truex et al., 2011). The desiccation process removes previously disposed and native water from the vadose zone and significantly decreases the water relative permeability of the desiccated region.

The performance of desiccation in terms of mitigating future moisture movement is related to the extent of moisture content reduction and the location and thickness of the desiccated zone (Truex et al., 2012). Thus, information concerning the distribution of moisture content reduction over time is needed

to monitor desiccation performance. Such data are important for determining when desiccation has met treatment goals and when the process can be stopped. Monitoring data can also be used to guide operational decisions, such as adjustments in system flow rates and injection gas properties. While nominal values for these injection parameters can be selected based on initial site characterization data, the impact of subsurface heterogeneities cannot be fully predicted and monitoring data are needed to assess the impact of these heterogeneities on desiccation performance. Several geophysical approaches for monitoring desiccation in terms of changes in the distribution of moisture content over time were investigated in a Hanford Site desiccation field test, including neutron moisture logging, ground penetrating radar (GPR), and time-lapse 3-D cross hole ERT.

The desiccation field test was conducted in the vadose zone at the Hanford Site 200 Area in the BC-Cribs and Trenches (BCCT). The total thickness of the vadose zone beneath the BCCT is about 100 m. Within the BCCT, about 110 million liters of aqueous LLW was infiltrated into multiple engineered cribs and trenches, primarily during the 1950s (DOE/RL, 1999). From 2004 to 2005, extensive electrical resistivity characterization efforts were completed across the BCCT site identifying large expanses of inorganic radioactive waste plumes within the vadose zone (Rucker and Fink, 2007; Rucker et al., 2004). The detailed mapping of the subsurface at BCCT contributed to its selection as a test site. The desiccation test consisted of a dipole dry gas injection/extraction well system where gas was passed through the target desiccation zone at a depth interval from ~9 to 15 m deep and a separation of ~12 m. Logging and ERT electrode boreholes were placed around the well to both characterize subsurface structure and monitor desiccation performance over time (Figure 19). Note that the subsurface electrical conductivity is governed in part by water content. Therefore, changes in water content are manifest as changes in conductivity within the time-lapse ERT images.

The baseline ERT imaging results are shown in Figure 20. The fine-grained lens with elevated bulk conductivity that straddles the injection well is the target desiccation zone. Such fine-grained lenses are known to have caused lateral contaminant flow in the Hanford vadose zone and generally exhibit elevated moisture content and contaminant levels when located beneath infiltration galleries. Zones with

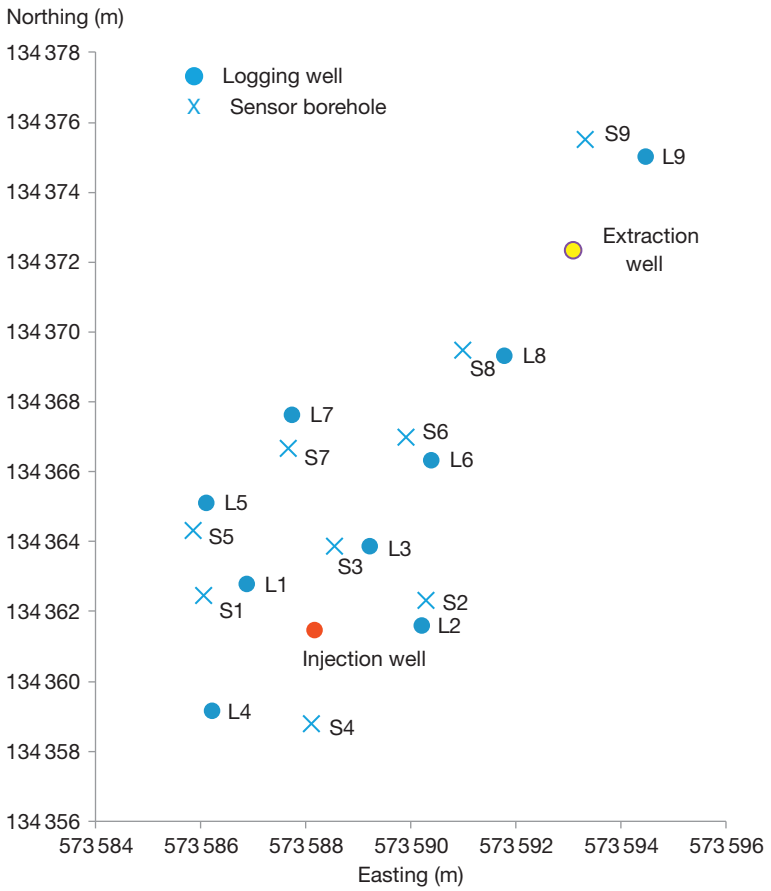


Figure 19 Plan view of desiccation field setup showing locations of gas injection and extraction wells within the Hanford BCCT. Blue dots mark locations of access boreholes for neutron moisture logging and cross hole GPR surveys. Blue crosses mark locations of vertical ERT electrodes arrays using for 3-D time-lapse ERT monitoring. In addition, a 30 m by 45 m gas-impermeable membrane barrier was installed at the surface centered over the well network to inhibit soil gas flow at the ground surface.

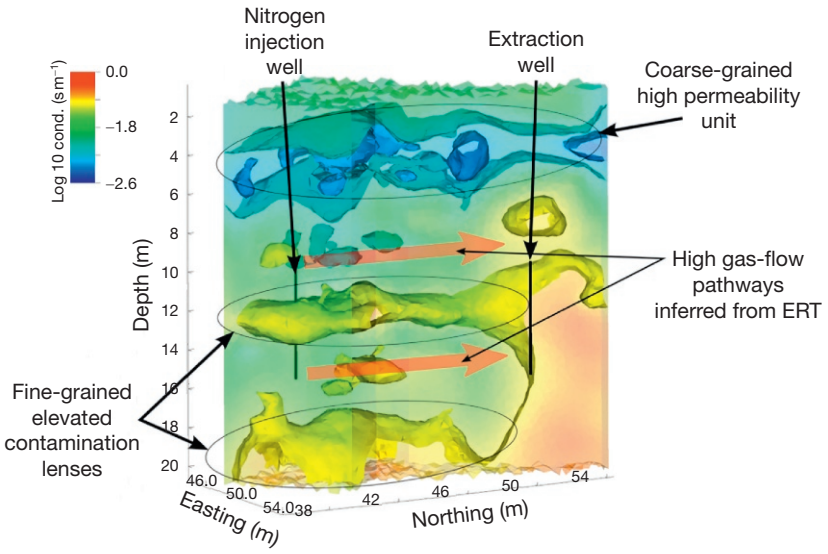


Figure 20 Predesiccation 3-D baseline ERT image showing bulk conductivity along selected isosurfaces. Higher conductivity lenses are diagnostic of fine-grained, elevated moisture units that tend to harbor contaminated pore water due to increase matric potential in comparison to coarser units. Lower conductivity lenses represent coarser-grained, higher permeability units that likely provide preferred gas flow pathways.

lower bulk conductivity are typically coarser-grained and more permeable and may provide preferred pathways for gas flow.

During the desiccation test, dry nitrogen gas was injected at a stable flow rate of 510 $m^3\ h^{-1}$ from 17 January 2011 to 30 June 2011 (164 days) except during a 13-day interval from 21 April 2011 to 4 May 2011 when there was no injection. Extraction of soil gas was maintained for the full test duration at a stable flow rate of 170 $m^3\ h^{-1}$. ERT data were collected prior to and during desiccation using 99 electrodes; 11 electrodes equally spaced from 6.25 to 21.5 m deep in each of the nine sensor wells. The electrical conductivities obtained from the inversions were used to estimate volumetric moisture content at each survey time using an approximation base on Archie's law (Archie, 1942; Truex et al., 2013). Core-scale testing on site sediments showed the electrical conductivity response to be primarily governed by decreases in water saturation as opposed to increases in fluid conductivity during desiccation, validating the assumption that fluid conductivity may be considered constant in time.

Figure 21 shows time-lapse 3-D changes in water content derived from the ERT images at selected times. The images indicate that much of the gas flow occurred below the fine-grained unit between 12 and 14 m deep shown in Figure 20 creating a primary desiccation zone just below the fine-grained lens. The figures also show the progression of the desiccation zone from the injection to the extraction well and the development of an upper desiccation zone near the top of the injection well. For validation, Figure 21 also shows a comparison of the ERT-derived changes in moisture content to those derived from interpolating neutron moisture logs collected 164 days after the start of injection. Although the relative volumes of the desiccated zones are somewhat different, both the ERT and neutron logs describe the same location and relative progression of the desiccated zone, regardless of the errors in each estimate (i.e., regularization smoothing and Archie's law assumptions for ERT, calibration, and interpolation errors from neutron logs). Overall, the ERT monitoring demonstrated the capability to remotely monitor 3-D changes in moisture content caused by desiccation. In addition, the approach is scalable to larger applications where well-based neutron monitoring becomes less reliable due to the larger distance between monitoring wells.

11.17.11 Underground Infrastructure Mapping and Site Clearance

Moving the HLW from the five separations plants to the individual tank farms and other waste disposal sites occurred through a series of overland, underground, and shallowly buried pipelines. Nearly 200 miles of abandoned pipelines traverse the Hanford Site and it is not uncommon for there to be over 10 miles of pipeline within a single tank farm (Jaraysi et al., 2007). The pipelines were mainly installed during original tank farm construction, dating back to the mid-1940s (Knepp, 2002). Pipeline attributes are diverse and complex, consisting of a variety of construction materials, differing diameters, and degree of pressurization (Jaraysi et al., 2007). In several cases, pipeline integrity has been compromised and pipe breaks, failed joints, plugging, and poorly functioning valves have led to unplanned releases (Knepp, 2002). A comprehensive investigation into the cause of the pipeline failures determined that most failures were caused by local conditions adjacent to the pipe (e.g., corrosion due to soil, water, and loss of coating), construction, or operating practices (Lambert, 2005). Despite adding anticorrosion measures, including cathodic protection systems, asphalt wrap, and

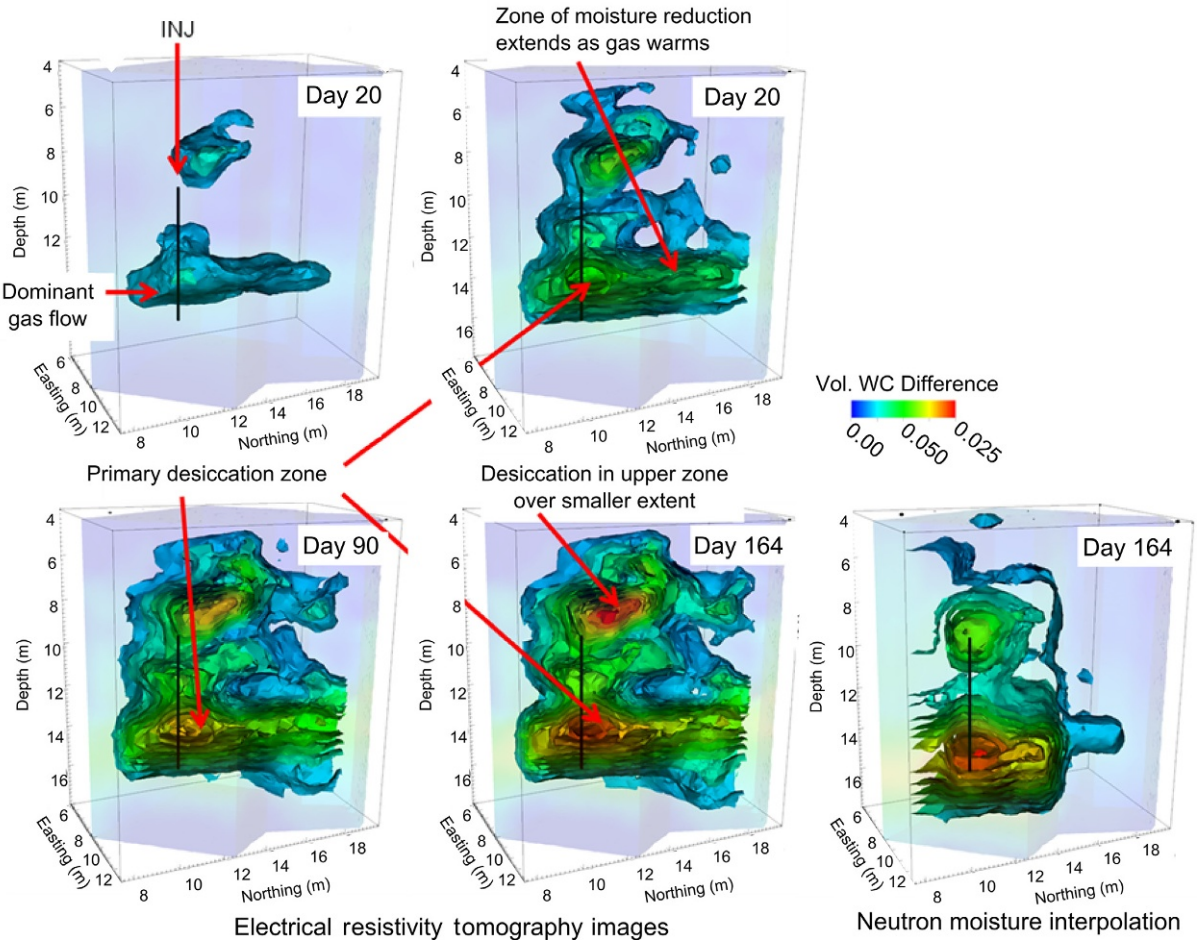


Figure 21 3-D view of the change in volumetric moisture content over time derived from cross hole electrical resistivity tomography and neutron moisture data. Reproduced from Truex MJ, Johnson TC, Strickland CE, Peterson JE, and Hubbard SS (2013) Monitoring vadose zone desiccation with geophysical methods. *Vadose Zone Journal* 12, vzj2012.0147.

eventually concrete encasement, pipe failure was still known to occur (Fink et al., 2011).

A key closure element for the Hanford Site is the development of remedial actions for the pipelines, which may incorporate various aspects of in-place stabilization (e.g., grouting) or complete removal (Jaraysi et al., 2007). In support of this effort, site characterization of the existing pipeline inventory has been extensively performed. Over the past decade, geophysical surveys with frequency-domain electromagnetic induction (EM), magnetic gradiometry, and GPR have been conducted in support of pipeline mapping, given that as-built drawings of engineered systems have been known to be inaccurate (Rucker, 2010). Within the tank farm, GPR (see **Chapter 11.07**) is the preferred method due to surface metallic objects interfering with EM and magnetic gradiometry interpretations. GPR survey lines are acquired at line spacings between 1.5 and 6 m, traversed in orthogonal directions with the support of heads up Global Positioning System (GPS) navigational aids to ensure complete coverage. For example, the TX and TY farms were surveyed with 30 line kilometers of GPR over an area 6.5 ha. From these data, diffraction hyperbolas were picked from the raw radargrams (i.e., stacked wiggle traces of voltage amplitude) and location and depth of these picks were tracked to develop maps of likely pipe locations. **Figure 22** shows an example of radargram hyperbola picks from the TY Farm investigation (Myers et al., 2008a), where the picks are placed on a plan map to show relative placement. The picks are also segregated by confidence based on the geophysicist's ability to determine the strength of a particular diffraction target relative to background reflections. Along with tank infrastructure, as-built pipeline locations are shown to help add perspective to the GPR picks with regard to what is thought to be known conditions. In several examples, there is an overlap between the known and GPR data sets. In other cases, there are pipes within the as-built drawings that were not imaged by GPR and there are obvious linear features in the GPR picks that represent likely pipes not in the drawings. Depending on the end use of the data, higher-density GPR data may be reacquired in a particular area to help develop a more detailed picture of the pipeline layout. Drilling for subsurface contaminant characterization would be one example where more GPR data would be warranted.

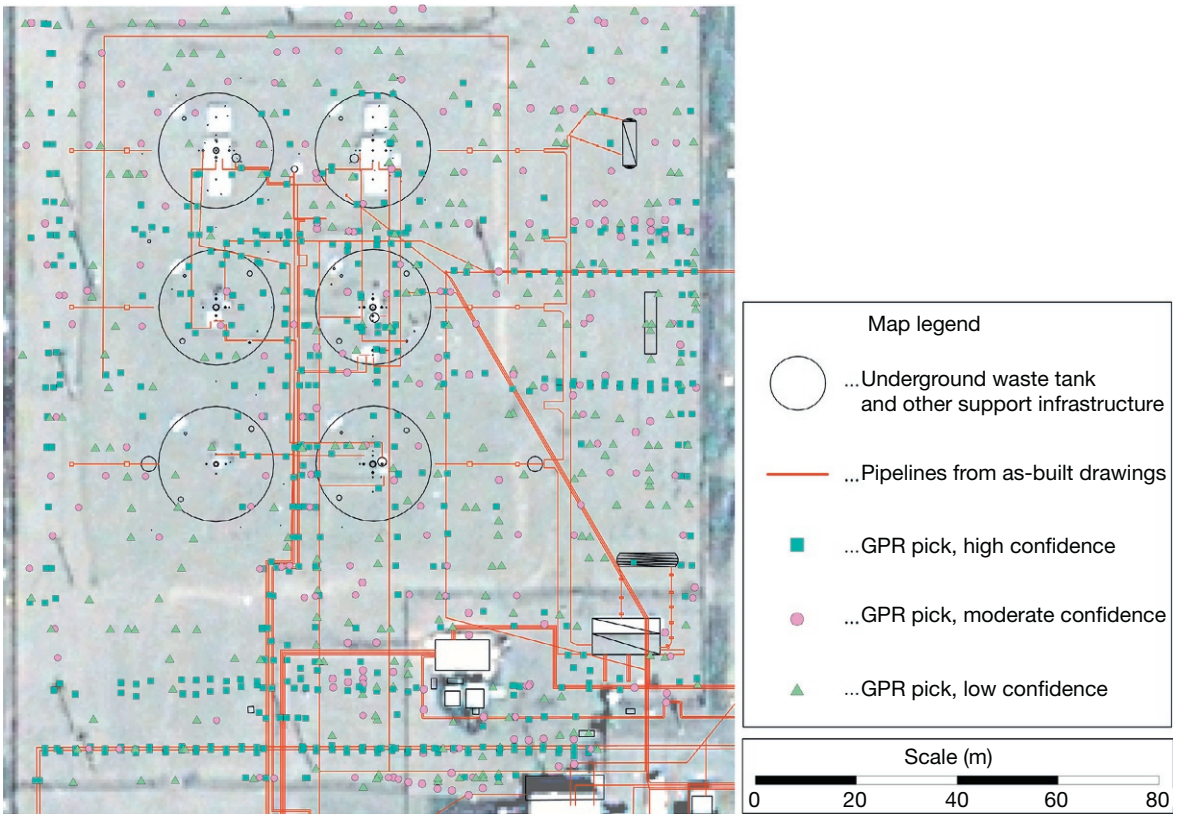


Figure 22 Map of TY tank farm showing the results of GPR hyperbola picks from radargrams. Picks are segregated into a confidence that the anomaly is from infrastructure, which may be influenced by moisture conditions, presence of cobbles, and depth.

Outside the tank farm, EM and magnetic gradiometry is preferred to GPR based on the ease of data acquisition and processing over broad areas. The instruments are fitted to a fiberglass cart and towed behind an all-terrain vehicle (ATV). The magnetic and EM data are usually acquired along parallel lines spaced ~2–3 m apart (Rucker, 2010). The speed of the ATV is controlled such that a nominal sampling rate of three measurements per meter is attained. Similar to the GPR setup, geographic positioning for the data is established using a navigation system mounted above the cart. Continuing with the same tank farm presented for the GPR example, Figure 23 shows EM data for the region outside TX and TY tank farms, covering ~40 ha (Myers et al., 2008b). Both in-phase and quadrature data were acquired at frequencies of 5, 10, and 20 kHz and these raw data were inverted for electrical conductivity and magnetic susceptibility. Figure 23 focuses on the 10 kHz data and shows high conductivity linear features traversing the site with a heavy corridor of piping to the east of TY tank farm. In addition to minimizing impact to pipes during drilling, these data are used to support electrical resistivity characterization and interpretation. Knowing the location of pipes and correlating to conductive features within the resistivity sections help to minimize false-positive interpretations related to a potential contaminant disposed to the ground.

11.17.12 Long-Term Legacy: The Role of Geophysics

The USDOE currently maintains one of the largest environmental cleanup programs in the world. At some sites, systems designed to monitor vadose zone, surface water, sediment, and groundwater contamination require millions of dollars per year to operate. In addition, at many of these sites, contaminants are expected to persist for decades to centuries, posing requirements and associated challenges for long-term monitoring. These challenges provide a significant scientific opportunity to leverage existing geophysical monitoring capabilities and to develop new capabilities for long-term monitoring challenges (Beam and Chamberlain, 2012). For example, the development of conceptual site models that incorporate the system as a whole will require remote sensing information to meet the data demand necessary to create such models. Flux-based contaminant assessment approaches and multiple lines of evidence that augment well-based sampling to improve subsurface interrogation will be necessary for long-term site management. Geophysical imaging capabilities that can inform site conceptual models and reduce monitoring costs while providing improved information will be critical for effective risk management at these sites. Because of this, near-surface geophysical techniques will continue to play a critical role in remediating and managing legacy subsurface contamination throughout the former USDOE nuclear weapons development complexes.

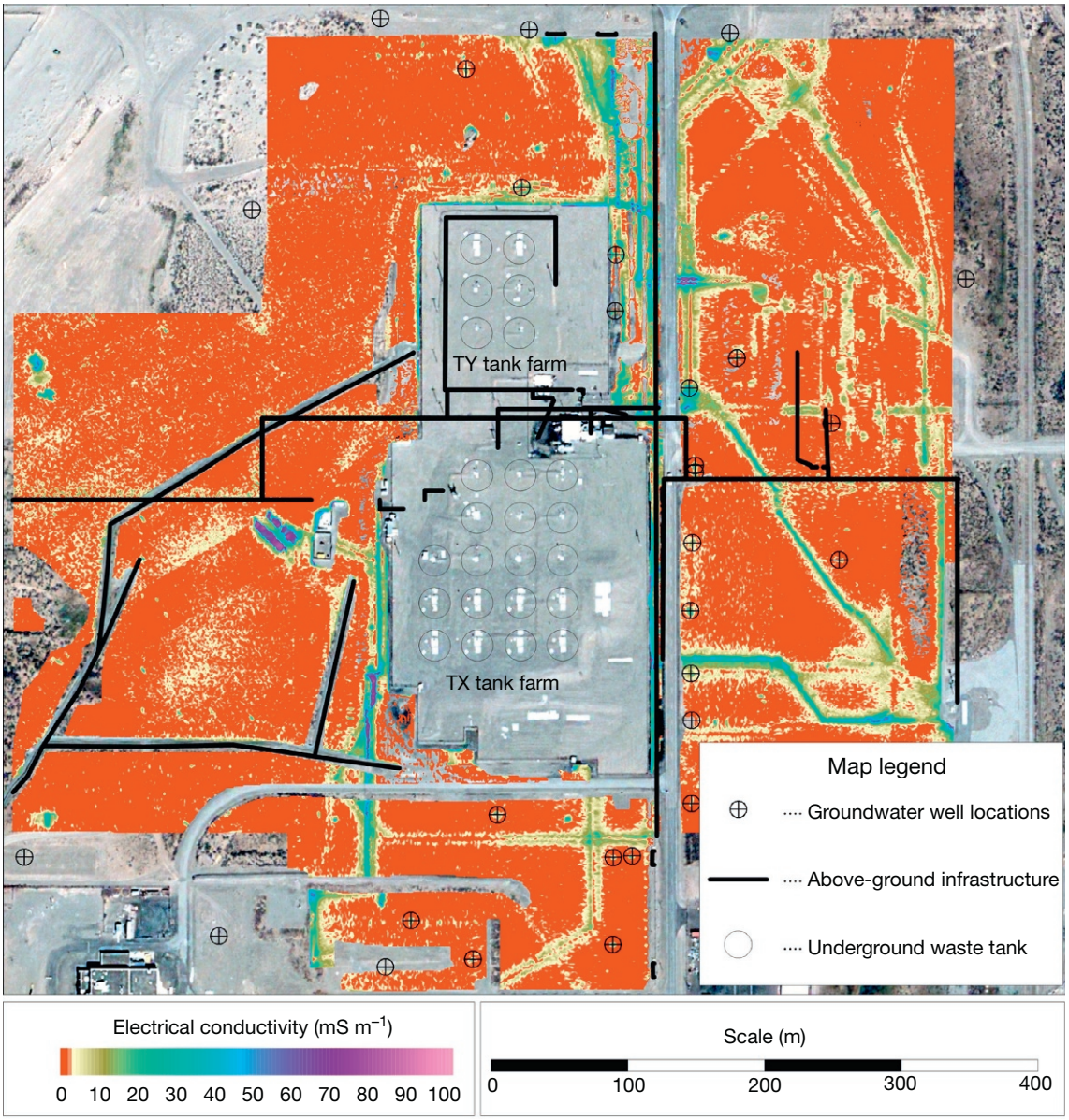


Figure 23 Near-surface electrical conductivity distribution outside of the TX and TY tank farms. The data were obtained using frequency-domain electromagnetic induction and are presented for the 10 kHz frequency. Linear features resulting from high conductivity values are a result of subsurface piping used to move waste from processing plants to the storage facility.

References

Archie GE (1942) The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the American Institute of Mining and Metallurgical Engineers* 146: 54–61.

Barnett DB, Gee GW and Sweeney MD (2002) Results of tank-leak detection demonstration using geophysical techniques at the Hanford Mock Tank Site – Fiscal year 2001. PNNL-13818. Richland, WA: Pacific Northwest National Laboratory. Available: http://www.pnl.gov/main/publications/external/technical_reports/PNNL-13818.pdf.

Barnett DB, Gee GW, Sweeney MD, et al. (2003) Results of performance evaluation testing of electrical leak-detection methods at the Hanford Site Mock Tank – FY 2002–2003. Richland, WA: Pacific Northwest National Laboratory. Available: http://www.pnl.gov/main/publications/external/technical_reports/PNNL-14192.pdf.

Beam P and Chamberlain GS (2012) *Scientific Opportunities for Monitoring at Environmental Remediation Sites [Online]*. Richland, WA: US Department of Energy, Office of Environmental Management. http://www.pnnl.gov/main/publications/external/technical_reports/PNNL-21379.pdf.

Bratton WL and Boger RM (2002) Ex-tank LDMM technology assessment and down-selection report. RPP-10604, Rev 1. Richland, WA: CH2M Hill Hanford Group, Inc.

Calendine S, Rucker DF, Fink JB and Levitt MT (2011) Automated leak detection of buried tanks using geophysical methods at the Hanford nuclear site. SAGEEP 2011. *Annual Meeting of the Environmental and Engineering Geophysical Society*, Charleston, SC, 10–14 April 2011.

Daily W, Ramirez A, and Binley A (2004) Remote monitoring of leaks in storage tanks using electrical resistance tomography: Application at the Hanford Site. *Journal of Environmental and Engineering Geophysics* 9: 11–24.

DOE/ORP (2005) *Initial Single-Shell Tank System Performance Assessment for the Hanford Site*. Richland, WA: US Department of Energy.

DOE/RL (1999) *Groundwater/Vadose Zone Integration Project Background Information and State of Knowledge*. Richland, WA: US Department of Energy.

DOE/RL (2008) *Deep Vadose Zone Treatability Test Plan for the Hanford Central Plateau*. Richland, WA: US Department of Energy, Richland Operations Office.

DOE/RL (2012) *Hanford Site Groundwater Monitoring for 2011*. Richland, WA: US Department of Energy, Richland Operations Office.

DOE/RL (2013) *Waste Treatment & Immobilization Plant Project [Online]*. Richland, WA: US Department of Energy. http://www.hanford.gov/page.cfm/WTP.

Dresel PE, Truex MJ, and Cantrell K (2008) *Remediation of Deep Vadose Zone Radionuclide and Metal Contamination: Status and Issues*. Richland, WA: Pacific Northwest National Laboratory.

Fink JB, Rucker DF, Levitt MT, and Skorska M (2011) Detection of historical pipeline leak plumes using non-intrusive, surface based geophysical techniques at the Hanford nuclear site. In: *Waste Management 2011*, Phoenix, AZ, 27 February to 3 March 2011.

Gephart RE (2003) *A Short History of Hanford Waste Generation, Storage, and Release*. Richland, WA: Pacific Northwest National Laboratory.

Gephart RE (2010) A short history of waste management at the Hanford Site. *Physics and Chemistry of the Earth* 35: 298–306.

Glaser DR (2013) Surface geophysical exploration comparisons with other data sources. WRPS-56461-VA. Richland, WA: Washington River Protection Solutions, LLC.

Glaser DR, Fink JB, Levitt MT, and Rucker DF (2008) A summary of recent geophysical investigations at the Department of Energy Hanford Nuclear Facility. Abstract #B-04. Presented at *2008 Chapman Conference on Biogeophysics*, 12–16 October. Portland, ME: AGU.

Hammond GE and Lichtner PC (2010) Field-scale model for the natural attenuation of uranium at the Hanford 300 Area using high-performance computing. *Water Resources Research* 46.

Hartman MJ, Morasch LF, and Webber WD (2006) Hanford site groundwater monitoring for fiscal year 2006. PNNL-16346. Richland, WA: Pacific Northwest National Laboratory.

Jaraysi MN, Badden JW, Lyon J, Quintero RA, Seeley P, and Lehman L (2007) Tank Farm Remediation Technology Development Project: An exercise in technical and regulatory collaboration. In: *Waste Management 2007*, Tucson, AZ, 25 February to 1 March 2007.

Johnson ME and Field JG (2010) Hanford SX-farm leak assessments report. RPP-ENV-39658. Richland, WA: Washington River Protection Solutions, LLC.

Johnson TC, Slater LD, Ntarlagiannis D, Day-Lewis FD, and Elwaseif M (2012) Monitoring groundwater-surface water interaction using time-series and time-frequency analysis of transient three-dimensional electrical resistivity changes. *Water Resources Research* 48.

Johnson TC, Versteeg RJ, Ward A, Day-Lewis FD, and Revil A (2010) Improved hydrogeophysical characterization and monitoring through parallel modeling and inversion of time-domain resistivity and induced-polarization data. *Geophysics* 75: Wa27–Wa41.

Johnson TC, Hammond GE, Versteeg RJ, and Zachara JM (2013) Real-time 4D ERT monitoring of river water intrusion into a former nuclear disposal site using a transient warping-mesh water table boundary. AGU 2013 Fall Meeting. San Francisco, CA: AGU.

Johnson TC and Wellman DM (2013) *Re-inversion of Surface Electrical Resistivity Tomography Data from the Hanford Site B-Complex [Online]*. Richland, WA: Pacific Northwest National Laboratory. http://www.pnnl.gov/publications/abstracts.asp?report=466710, Accessed PNNL-22520.

Johnson TC, Wellman DM, and Triplett M (2014) Modeling of electrical resistivity data in the presence of electrically conductive well casings and waste storage tanks. In: *Waste Management Symposia*. Phoenix, AZ: Waste Management.

Key RT (1977) Nuclear waste tank and pipeline external leak detection systems. ARH-ST-127. Richland, WA: Atlantic Richfield Hanford Company.

Knepp AJ (2002) Field investigation report for waste management area S-SX. RPP-7884. Richland, WA: CH2M HILL Hanford Group, Inc.

Lambert SL (2005) Residual waste inventories in the plugged and abandoned pipelines at the Hanford Site. RPP-25133. Richland, WA: CH2MHill Hanford Group.

Lindenmeier CW, Lindberg MJ, Baum SR, et al. (2002) *Characterization of Vadose Zone Sediment: Borehole C3103 Located in the 216-B-7A Crib and Selected Samples from Borehole C3104 Located in the 216-B-38 Trench Near the BX Tank Farm*. Richland, WA: Pacific Northwest National Laboratory.

Myers DA, Cubbage B, Brauchla R, O'brien G, Bergeron M, and Henderson C (2008a) Surface geophysical exploration of TX-TY tank farms at the Hanford Site: Results of background characterization with ground penetrating radar. RPP-RPT-38104. Richland, WA: CH2M HILL Hanford Group, Inc.

Myers DA, Rucker DF, Levitt M, Cubbage B, and Henderson C (2008b) Surface geophysical exploration of TX-TY tank farms at the Hanford Site results of background characterization with magnetics and electromagnetics. RPP-RPT-36893. Richland, WA: CH2M HILL Hanford Group, Inc.

NRC (2000) *Research Needs in Subsurface Science – U.S. Department of Energy's Environmental Management Science Program*. Washington, DC: National Academies Press.

Oostrom M, Freedman VL, Wietsma TW, Dane JH, and Truex MJ (2012) Effects of porous medium heterogeneity on vadose zone desiccation: Intermediate-scale laboratory experiments and simulations. *Vadose Zone Journal* 11.

Oostrom M, Wietsma TW, Dane JH, Truex MJ, and Ward AL (2009) Desiccation of unsaturated porous media: Intermediate-scale experiments and numerical simulation. *Vadose Zone Journal* 8: 643–650.

Oostrom M, Wietsma TW, Strickland CE, Freedman VL, and Truex MJ (2012) Sensor and numerical simulator evaluation for porous medium desiccation and rewetting at the intermediate laboratory scale. *Vadose Zone Journal* 11.

Peterson SW (2010) *Measurement of Compressional-Wave Seismic Velocities in 29 Wells at the Hanford Site*. Richland, WA: CH2M-Hill.

Peterson RE, Rockhold ML, Serne RJ, Thorne PD, and Williams MD (2008) *Uranium Contamination in the Subsurface Beneath the 300 Area, Hanford Site, Washington*. Richland, WA: Pacific Northwest National Laboratory.

Ramirez A, Daily W, Binley A, LaBrecque D, and Roelant D (1996) Detection of leaks in underground storage tanks using electrical resistance methods. *Journal of Environmental and Engineering Geophysics* 1: 189–203.

Reich FR (2004) *SST Deployment Demonstration and Injection Leak Testing of the HRR Long Electrode LDM System*. Richland, WA: COGEMA-Engineering, Inc.

Rodgers MJ (2013) Waste tank summary report for month ending August 31, 2013. HNF-EP-0182, Rev. 305. Richland, WA: Washington River Protection Solutions, LLC.

Rucker DF and Fink JB (2007) Inorganic plume delineation using surface high-resolution electrical resistivity at the BC cribs and trenches site, Hanford. *Vadose Zone Journal* 6: 946–958.

Rucker D, Levitt M, Bergeron M, Greenwood J, O'Brien G, McNeill M, and Henderson C (2007) *Surface Geophysical Exploration of the B, BX, and BY Tank Farms at the Hanford Site*. RPP-34690, Rev 0, Richland, WA: CH2M HILL.

Rucker DF (2012) Enhanced resolution for long electrode ERT. *Geophysical Journal International* 191: 101–111.

Rucker DF, Crook N, Glaser D, and Loke MH (2012) Pilot-Scale field validation of the long electrode electrical resistivity tomography method. *Geophysical Prospecting* 60(6): 1150–1166.

Rucker DF, Fink JB, Glaser DR, Gee GW, and Sweeny MK (2004) Comprehensive geophysical investigation over a former radioactive waste site, Hanford, Washington [abstract]. *Eos (American Geophysical Union Transactions)* 85(47): F772–F773. http://adsabs.harvard.edu/abs/2004AGUFM.H23A1115R.

Rucker DF, Fink JB, and Loke MH (2011) Environmental monitoring of leaks using time-lapsed long electrode electrical resistivity. *Journal of Applied Geophysics* 74: 242–254.

Rucker DF, Loke MH, Levitt MT, and Noonan GE (2010) Electrical-resistivity characterization of an industrial site using long electrodes. *Geophysics* 75(4): WA95–WA104.

Rucker DF, Myers DA, Cubbage BD, et al. (2013) Surface geophysical exploration: Developing noninvasive tools to monitor past leaks around Hanford's tank farms. *Environmental Monitoring and Assessment* 185: 995–1010.

Schofield JS (2006) Tank 241-S-102 high resolution resistivity leak detection and monitoring test report. RPP-30121. Rev Richland, WA: CH2M Hill Hanford Group, Inc.

Slater LD, Ntarlagiannis D, Day-Lewis FD, et al. (2010) Use of electrical imaging and distributed temperature sensing methods to characterize surface water-groundwater exchange regulating uranium transport at the Hanford 300 Area, Washington. *Water Resources Research* 46.

Truex MJ, Johnson TC, Strickland CE, Peterson JE, and Hubbard SS (2013) Monitoring vadose zone desiccation with geophysical methods. *Vadose Zone Journal* 12, vzj2012.0147.

Truex MJ, Oostrom M, Freedman VL, et al. (2011) *Laboratory and Modeling Evaluations in Support of Field Testing for Desiccation at the Hanford Site*. Richland, WA: Pacific Northwest National Laboratory.

Truex MJ, Oostrom M, Strickland CE, Chronister GB, Benecke MW, and Johnson CD (2012) Field-scale assessment of desiccation implementation for deep vadose zone contaminants. *Vadose Zone Journal* 11.

US Environmental Protections Agency (1996) EPA Superfund Record of Decision: Hanford 300-Area (USDOE) OU 01, 02 Benton County.

Wallin EL, Johnson TC, Greenwood WJ, and Zachara JM (2013) Imaging high stage river-water intrusion into a contaminated aquifer along a major river corridor using 2-D time-lapse surface electrical resistivity tomography. *Water Resources Research* 49(3): 1693–1708.

WHC (1993) *Sampling and Analysis of 300-FF-5 Operable Unit Springs and Near-Shore Sediments and River Water*. Richland, WA: Westinghouse Hanford Company.

Williams JC (1996) *Liquid Radioactive Waste Discharges from B Plant to Cribs*. Richland, WA: Westinghouse Hanford Co.

Williams MD, Rockhold ML, Thorne PD, and Chen Y (2008) *Three-Dimensional Groundwater Models of the 300 Area at the Hanford Site, Washington State, Report*. Richland, WA: Pacific Northwest National Laboratory.

Williams BA, Thompson MD, and Miller SF (2012a) *Integrated Surface Geophysical Investigation Results at Liquid Effluent Retention Facility, 200 East Area, Hanford, Washington*. Richland, WA: CH2MHill Plateau Remediation Company.

Williams BA, Thompson MD, and Miller SF (2012b) *Interpretation and Integration of Seismic Data in the Gable Gap*. Richland, WA: CH2MHill Plateau Remediation Company.

Williams BA, Thompson MD, and Miller SF (2012c) *Land streamer and Gimbaled Geophones Phase II-200 Areas: High-Resolution Seismic Reflection Survey at the Hanford Site*. Richland, WA: CH2MHill Plateau Remediation Company.

Williams BA, Thompson MD, and Miller SF (2012d) *Seismic Reflection Investigation at the Liquid Effluent Retention Facility, 200 East Area, Hanford Site Richland, Washington*. Richland, WA: CH2MHill Plateau Remediation Company.